AMERICAN
JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

PUBLISHED UNDER THE JOINT AUSPICES OF
THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LXV, Number 2
APRIL, 1943

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

CONTENTS

The American Journal of Mathematics will appear four times yearly.

The subscription price of the Journal for the current volume is $7.50 (foreign postage 50 cents); single numbers $2.00.

A few complete sets of the Journal remain on sale.

Papers intended for publication in the Journal may be sent to any of the Editors.

Editorial communications may be sent to Professor F. D. Murnaghan at The Johns Hopkins University.

Subscriptions to the Journal and all business communications should be sent to The Johns Hopkins Press, Baltimore, Maryland, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

FORMAL REDUCTIONS OF THE GENERAL COMBINATORIAL DECISION PROBLEM.*

By EMIL L. POST.

1. Introduction. It is not new to the literature that the usual form of a symbolic logic with its parenthesis notation and infinite set of variables can be transformed into one in which the *enunciations,* i. e., formulas of the system, are finite sequences of letters,[1] the different letters constituting a once-and-for-all given finite set. If the primitive letters of such a system are represented by $a_1, a_2, \cdots, a_\mu$, an arbitrary enunciation of the system will take the form $a_{i_1} a_{i_2} \cdots a_{i_n}$, $n = 1, 2, 3, \cdots$, $i_j = 1, 2, \cdots, \mu$. In describing the basis of such a system it is convenient to use new letters to represent finite sequences of the above primitive letters. If then $A, B, \cdots, E$ represent the sequences $a_{i_1} a_{i_2} \cdots a_{i_\rho}$, $a_{j_1} a_{j_2} \cdots a_{j_\sigma}, \cdots, a_{m_1} a_{m_2} \cdots a_{m_\phi}$ respectively, $AB \cdots E$ will represent the sequence $a_{i_1} a_{i_2} \cdots a_{i_\rho} a_{j_1} a_{j_2} \cdots a_{j_\sigma} \cdots a_{m_1} a_{m_2} \cdots a_{m_\phi}$.

We shall say that such a system is in *canonical form* if its basis has the following structure.[2] The *primitive assertions* of the system are a specified finite set of enunciations of the above form. The operations of the system are a specified finite set of *productions,* each of the following form:

$$g_{11} P_{i'_1} g_{12} P_{i'_2} \cdots g_{1m_1} P_{i'_{m_1}} g_{1(m_1+1)}$$
$$g_{21} P_{i''_1} g_{22} P_{i''_2} \cdots g_{2m_2} P_{i''_{m_2}} g_{2(m_2+1)}$$
$$\cdots \cdots \cdots \cdots$$
$$g_{k_1} P_{i_1}^{(k)} g_{k_2} P_{i_2}^{(k)} \cdots g_{km_k} P_{i_{m_k}}^{(k)} g_{k(m_k+1)}$$
produce
$$g_1 P_{i_1} g_2 P_{i_2} \cdots g_m P_{i_m} g_{m+1}.$$

---

* Received November 14, 1941; Revised April 11, 1942.

[1] More exactly, "strings" of "marks," to use terms of C. I. Lewis (*A Survey of Symbolic Logic,* Berkeley, 1918: chapter VI, sec. III).

[2] This formulation stems from the "Generalization by Postulation" of the writer's "Introduction to a general theory of elementary propositions," *American Journal of Mathematics,* vol. 43 (1921), pp. 163-185 (see p. 176). We take this opportunity to make the following *Emendation*: Lemma 1 thereof (pp. 177-178) requires the added condition that the expressions replacing the r's do not involve any letter upon which a substitution is made in the given deductive process. This necessitates several minor changes in the proof of the theorem there following. Actually, both Lemma 1 and its companion Lemma 2 admit of further simplification, with the proof of the theorem then being valid as it stands.

In this display the g's represent specified sequences of the primitive a's, including the null sequence, while the P's represent the operational variables of the production, and, in the application of the production, may be identified with arbitrary sequences of this type. In this notation, the distinct operational variables of a given production are to constitute the finite set of symbols $P_1, P_2, \cdots, P_M$ for some positive integral M. We then add the restriction that each operational variable in the conclusion of a production is present in at least one premise of that production, it having been understood that each premise and conclusion has at least one operational variable. We further assume that no identification of the operational variables is permitted which would lead to the conclusion being null.[3] The *assertions* of the system are then the primitive assertions, and all enunciations obtainable by the repeated application of the given productions starting with the primitive assertions.[4] More precisely, the class of assertions is the smallest subclass of the class of enunciations which contains the primitive assertions, and which, for each admissible assignment of values to the operational variables of each of the given productions, contains the enunciation represented by the conclusion of

---

[3] For the proof of Section **2** to be universally valid it is necessary that the operations themselves exclude the possibility of a null assertion. Since a null conclusion could arise only from an operation whose conclusion consists of operational variables only, while under our restriction at least one of these operational variables is not to be null, we can achieve the desired automatic exclusion of the null assertion by replacing each such operation by the following equivalent finite set of operations. For each operational variable P_i in the conclusion of such an operation, and each primitive letter a_j form the operation obtained by replacing P_i by $a_j P_i$ throughout the given operation. This modification need only be made on the given system in canonical form; for the productions introduced in Section **2** all have their single premises consist of more than just operational variables, so that the nonexclusion of the null assertion would have no further effect on the system.

[4] That this leads to the constructive generation of the class of assertions is readily verified. In particular, in trying to identify the premises of a production with corresponding previously obtained assertions, an explicit hypothesis on the "rank" of the operational variables involved, rank of a sequence being the total number of letters therein, would immediately lead to their unique determination or impossibility of realization, and hence correspondingly to a unique conclusion or impossibility of derivation. Since the sum of the ranks of the fixed g's and the fixed number of operational variables of a premise must equal the fixed rank of the assertion that premise is to be identified with, only a finite number of such hypotheses are admissible, and all can be uniformly tried out. In practice, a system in canonical form will usually be so constructed that a given assertion can be written in the form of a given premise in one and only one way, if at all. This uniqueness is automatically achieved by systems using the parenthesis notation, and is, of course, obviously attained in the systems in normal form about to be mentioned.

the production whenever it contains the enunciations represented by the several premises of the production.

A very special case of the canonical form is what we term the normal form. A system in canonical form will be said to be in *normal form* if it has but one primitive assertion, and, each of its productions is in the form

$$g\,P$$

produces

$$P\,g'.$$

The main purpose of the present paper is to demonstrate that every system in canonical form can formally be reduced to a system in normal form. The two forms may therefore in fact be said to be *equipotent*. More precisely, we prove the following

THEOREM. *Given a system in canonical form with primitive letters* $a_1, a_2, \cdots, a_\mu$, *a system in normal form with primitive letters* $a_1, a_2, \cdots, a_\mu$, $a'_1, a'_2, \cdots, a'_\mu$ *can be set up such that the assertions of the system in canonical form are exactly those assertions of the system in normal form which involve no other letters than* $a_1, a_2, \cdots, a_\mu$.

As a result of this theorem the decision problem for a system in canonical form is reduced to the decision problem for the corresponding system in normal form. For an enunciation of the former system is an assertion when and only when it is an assertion of the latter system. Hence any procedure which could effectively determine for an arbitrary enunciation of the system in normal form whether it is or is not an assertion thereof would automatically do the same for the system in canonical form. Now by methods such as those referred to in the opening sentence of this introduction, it can be shown that the problem of determining for an arbitrary well-formed formula in the λ-calculus of Church whether it has or has not a normal form (Church) [5] can be reduced to the decision problem for a particular system in our canonical form. While Church has proved the above problem unsolvable in a certain technical sense, in the interest of economy we invoke his identification of λ-definability with effective calculability to conclude that as a result the decision problem for that particular system in canonical form, and hence for the class of systems in canonical form, is unsolvable. We are thus led to the more surprising result

[5] Alonzo Church, "An unsolvable problem of elementary number theory," *American Journal of Mathematics*, vol. 58 (1936), pp. 345-363.

that *there can be no effective procedure for determining for an arbitrary system in normal form and arbitrary enunciation thereof whether that enunciation is or is not an assertion of the system.* That is, the decision problem is unsolvable for the class of normal systems, and indeed, by the previous argument, for a certain particular one of them.[6]

The present paper is not the place to review the reasons why the equivalent mathematical definitions of combinatory solvability based on the technical concepts of λ-definability, general recursive function, and computability [7] can confidently be accepted as being the complete equivalent of combinatory solvability in the intuitive sense. Granting the initial establishment of the unsolvability of a particular decision problem by virtue of its being directly coextensive with the technical definition of solvability adopted, the chief method of establishing the unsolvability of further removed decision problems is by reducing the known unsolvable problem, by more or less ingenious formal devices, to those other problems.[8] Our reduction of the decision problem for the complicated canonical form to that of the simple normal form illustrates this in some measure. And it may be that because of its formal simplicity, the normal form may lend itself more readily to representation in specialized mathematical developments, and the unsolvability of its decision problem thus lead to the unsolvability of various hitherto unsolved decision problems of classical mathematics.

Of more immediate promise is the fact that the concepts of the present paper, with the help of its basic theorem, easily lead to an independent approach to unsolvable problems which may be far simpler than, say, the λ-calculus of Church. In this connection we may note that if we define a *normal set* of

---

[6] Absolutely unsolvable, that is, to use a phrase due to Church. By contrast, the undecidable propositions of Gödel's epoch making paper of 1931 (see footnote 7) are but relatively undecidable, the very proof of their undecidability in the given logic leading to an extension of that logic in which they are, indeed, proved to be true. A fundamental problem is the question of the existence of absolutely undecidable propositions, that is, propositions which in some *a priori* fashion can be said to have a determined truth-value, and yet cannot be proved or disproved by any valid logic.

[7] For the first two see the paper referred to in footnote 5, for the third see A. M. Turing, "On computable numbers, with an application to the Entscheidungsproblem," *Proceedings of the London Mathematical Society* (2), vol. 42 (1937), pp. 230-265. We might also add the writer's "Finite combinatory processes-formulation I," *Journal of Symbolic Logic*, vol. 1 (1936), pp. 103-105. The basic paper is, of course, that of Kurt Gödel, "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I," *Monatshefte für Mathematik und Physik*, vol. 38 (1931), pp. 173-198.

[8] A very important instance of such a reduction is Gödel's transformation of the iterative recursive proposition into the non-iterative arithmetical proposition.

sequences on $a_1, a_2, \cdots, a_\mu$ as the set of assertions on those letters only of any system in normal form with primitive letters $a_1, a_2, \cdots, a_\mu$ and a finite number of additional letters, and a *canonical set* similarly via a system in canonical form, then the above theorem has as an immediate consequence the

COROLLARY. *The class of canonical sets is identical with the class of normal sets.*

For every normal set is *ipso facto* a canonical set; while if a canonical set is given by a certain system in canonical form, the theorem shows that it is also given by the corresponding system in normal form, and hence is a normal set. Now the canonical form naturally lends itself to the generating of sets by the method of definition by induction, while redefining the resulting canonical sets as normal sets makes it easy to use them as building blocks in further constructions. As a result of this alternating use of the canonical form as method, normal set as object, the Church development is easily paralleled.[9] And since at each step only normal sets of sequences are obtained, we are led to identify the intuitive concept of *generated set* with normal set for much the same reasons that led Church to identify effective calculability with λ-definability. Under this identification, the intuitive concept of a *solvable set* of sequences on $a_1, a_2, \cdots, a_\mu$, i. e., one for which there is an effective procedure for determining whether a given sequence on those letters is or is not in the set, becomes precisely the *binormal set,* i. e., a set such that both it, and its complement with respect to the set of all finite sequences on $a_1, a_2, \cdots, a_\mu$, are normal. The resulting definition of solvability then easily leads to the unsolvability of the decision problem for the class of normal systems, as well as for a particular one of them. We may note this interchange of primary and secondary concept as compared with the Church development; for normal set corresponds to recursively enumerable set, binormal set to (general) recursive set.[10]

---

[9] More completely, the Gödel, Church, Kleene, Rosser development.

[10] While this equivalence undoubtedly follows from the reduction of the λ-calculus to a system in normal form, it would probably be more easily established by way of Turing's concept of computability. A few initial properties of normal and binormal sets may here be noted. With the ordinary Boolean operations on classes in question, the class of all normal sets constitutes a (distributive) lattice, of all binormal sets, a Boolean ring, of all binormal sets on a given finite set of letters, a Boolean algebra. Every infinite normal set contains an infinite binormal set. Query: Does there exist an infinite set which is the complement of a normal set, relative to the given set of letters, and does not contain an infinite binormal set? [Added in proof: yes]. There is no theoretical loss of generality in restricting ourselves to normal sets on a single letter

Before turning to the proof of our basic theorem given in the next section, we wish to mention a further transformation of the normal form which is of interest for its juxtaposition of the solvable and unsolvable, and state a problem which largely determined the direction taken by the reductions of the next section, and may offer further opportunities for unsolvability proofs. By making the question of whether a given sequence has a certain succession of primitive letters at one end depend on a related sequence having a corresponding succession of letters at the other end, the following result can be proved. Given any system in normal form on primitive letters $a_1, a_2, \cdots, a_\mu$, an enunciation P thereof will be an assertion when and only when $\alpha P \alpha$ is an assertion in a corresponding effectively derivable system in canonical form on letters $a_1, a_2, \cdots, a_\mu, \alpha, a'_1, a'_2, \cdots, a'_\mu$, having a finite number of primitive assertions, and a finite number of operations of the following forms,

gP	Pg	g_1P	Pg_1
produces	*produces*	g_2P	Pg_2
$g'P$,	Pg' ,	*produce*	*produce*
		$g'P$,	Pg'.

A solution of the decision problem for the derived system then immediately yields a solution of the decision problem for the given system. Much of the simplicity of the normal form is thus given up in order to have the g' of the conclusion of each production on the same side of the operational variable as the g, or now g's, of the premise, or premises. But it is this very fact that leads to a solution of the decision problem for certain classes of these systems given *ab initio*. Indeed, for those systems in which all the productions have the g's on the same side a solution of the decision problem follows almost immediately from the solution of a decision problem given by the writer in a former paper.[11] This solution has been extended by the writer to those of the above systems having only first order productions, i. e., productions with but one premise, and it has at least been seen by the writer how to extend these

---

a—normal sets of natural numbers essentially. The following is then probably the analogue for normal systems, i. e., systems in normal form, of Turing's universal computing machine. A fixed finite set of normal operations involving a and a single additional letter b can be set up such that by varying a single primitive assertion on a and b all normal sets on a are obtained.

[11] "On a simple class of deductive systems," abstract, *Bulletin of the American Mathematical Society*, vol. 27 (1921), pp. 396-7. The systems in question are those systems of the formulation referred to in footnote 2 whose primitive functions are all functions of one argument.

solutions to those of the above systems in which only the second order productions are restricted to having their g's all on the same side, e. g., to systems having only productions of the first three of the above four types. For the class of all of the above systems, as indeed for a certain particular one of them, the decision problem is of course unsolvable as a consequence of the corresponding result for systems in normal form.

The problem referred to above takes two related forms. Both forms employ the following "tag" operations as we shall call them. Given a positive integer ν, and μ symbols which may be taken to be $0, 1, \cdots, \mu - 1$, we associate with each of the μ symbols a finite sequence of these symbols as follows.

$$0 \rightarrow a_{0,1}\, a_{0,2} \cdots a_{0,\nu_0}$$

$$1 \rightarrow a_{1,1}\, a_{1,2} \cdots a_{1,\nu_1}$$

$$\cdots\cdots\cdots\cdots$$

$$\mu - 1 \rightarrow a_{\mu-1,1}\, a_{\mu-1,2} \cdots a_{\mu-1,\nu_{\mu-1}}.$$

It is understood that in each sequence the same symbol may occur several times, and that a particular associated sequence may be null. Given any non-null sequence

$$B = b_1\, b_2 \cdots b_l$$

on the symbols $0, 1, \cdots, \mu - 1$, a unique derived sequence B' on those symbols is determined as follows. To the right end of B adjoin the sequence associated with b_1, the first symbol of B, and from the left end of the resulting augmented sequence remove the first ν symbols—all if there be less than ν symbols. Starting with a given tag operation, and a given sequence A on its primitive symbols, we can then iterate the tag operation to yield $A_1 = A, A_2 = A'_1$, $A_3 = A'_2, \cdots$, the process terminating when and only when the null sequence is thus obtained. The first form of the problem of "tag" for a given tag operation is then to find an effective procedure for determining of an arbitrary initial sequence whether the above iterative process does or does not terminate.[12] In the second form of the problem we assume both the tag operation and the

[12] In an earlier formulation of this problem we merely checked off each successive ν-th letter of the sequence starting with b_1, at the same time adding the corresponding associated sequences to the right end of the sequence. Whether the iterative process terminated or not then depended on whether the constantly advancing check mark did or did not overtake the monotonically advancing right end of the sequence, whence the suggestive name of "tag" given the problem by B. P. Gill.

initial sequence to be given, and ask for an effective procedure for determining of an arbitrary sequence whether it is or is not one of the sequences obtained from the given sequence by the iteration of the given tag operation.

The first form of the problem of tag was intensely studied by the writer.[13] Extended by the further dichotomy of the non-terminating cases to the periodic (sequences bounded), and divergent (sequences unbounded), the problem was completely solved for all cases in which both μ and ν are 2. But little real progress can be reported for μ or ν greater than 2, the problem for such a simple basis as $0 \rightarrow 00$, $1 \rightarrow 1101$, $\nu = 3$ having proved intractable.[14] In its second form the problem is almost the decision problem for a special type of normal system. In fact, we may define a *monogenic normal system* as one in which the g's of the premises form a *complete set,* i. e., a set $g_1, g_2, \cdots, g_k$ such that each of the sequences of length equal to the maximum length ν of the g's can be written in the form $g_i P$ for one and only one i. Except for the tag operation being applicable to sequences of length less than ν, a system of tag in its second form is then a monogenic normal system in which the g's constitute all of the μ^ν sequences in question, while the corresponding g''s are identical for all g's having the same initial symbol.

For a given tag operation the solution of the first form of the problem of tag probably leads to the solution of the second form of the problem. This is immediately so for those initial sequences which lead to termination or periodicity; and, while the mere hypothesis of divergence seems insufficient to guarantee a corresponding solution, the actual proof of divergence would probably make the definition of divergence effective, in which case the solution of the second form of the problem would again follow. For the writer, the little progress made in the solution of the first form of the problem make both forms, in their full generality, candidates for unsolvability proofs. Even more so, therefore, the decision problem for the class of monogenic normal systems. Among normal systems there is a "complete normal system" to which every normal system can be reduced, and whose decision problem is consequently unsolvable. A most interesting situation would obtain should it be shown that the complete normal system cannot be reduced to a monogenic normal system,

---

[13] During the writer's tenure of a Procter fellowship at Princeton University, 1920-1921.

[14] Numerous initial sequences actually tried led in each case to termination or periodicity, usually the latter.

while the decision problem for the class of monogenic normal systems is otherwise shown to be unsolvable.[15]

2. Reduction of the canonical form to the normal form. Our reduction of the canonical form to the normal form is the result of four successive reductions.[16] Each of these reductions yields a formulation which is included in the preceding formulation, but eliminates some formal complexities allowed in that preceding formulation. For a given system this simplification is achieved at the expense of an increase in the number of primitive letters employed, and in the number of productions appearing in its bases.

Our first reduction of an arbitrary system in canonical form is to one in which there is but one primitive assertion, and in which each production involves but a single premise. That one premise, and corresponding conclusion, however, may have all the complexity allowed for in the general canonical form. The general plan of the method involved is to formally introduce the logical products of arbitrary assertions of the given system, and operate within such products.

Let then S_1 be a system in canonical form with primitive letters $a_1, a_2, \cdots, a_\mu$, S_2 the system, about to be described, to which S_1 is to be reduced. With $a_1, a_2, \cdots, a_\mu$ also primitive letters of S_2, introduce two new primitive letters u and a_0 in S_2. When the logical product of assertions, $P_1, P_2, P_3, \cdots, P_n$ of S_1 is asserted in S_2, it will appear in the form

$$ua_0P_1\,a_0uua_0P_2\,a_0uuua_0P_3\,a_0\ldots\underbrace{u\ldots u}_{n}\,a_0P_na_0\,\underbrace{u\ldots u}_{n+1},$$

each P being flanked on either side by a_0. The separating u sequences are thus made to increase left to right by one each to enable us by the mere form of a premise to insure that certain operational variables therein must represent assertions of S_1, if that premise is to be identified with an assertion in S_2. The final basis for S_2 will reveal the necessary source of that insurance, i. e., that the only assertions of S_2 involving u are those of the above form. We shall call such an expression a product, the P's therein the factors of the product.

---

[15] It is easy to talk of obtaining a property of all normal solutions which could not be satisfied by a solution of a given decision problem; but this is probably equivalent to finding one of those not immediately obvious effectively calculable invariants of conversion which Church reports as still unfound in 1936. (See p. 358 of the paper referred to in footnote 5).

[16] Not counting a minor reduction needed to validate the last of the four.

We first introduce in the basis of S_2 certain productions whereby from the assertion of a product may be obtained the assertion of all products obtainable from the given product by a mere permutation of its factors. It suffices to allow for the interchange of any two consecutive factors. For the first two factors of a product this is achieved by

$$ua_0P_1a_0uua_0P_2a_0uuua_0S \quad produces \quad ua_0P_2a_0uua_0P_1a_0uuua_0S,$$

our system being so devised that each product appearing therein has at least three factors. This allows that last a_0 to be assumed. The u, uu, uuu of the premise are then "maximal" u sequences. As these u sequences differ by one each, P_1 and P_2 must be free from u's, and hence, by our induction, be the two initial factors of the product. The interchange then results via the production. For two consecutive factors neither starting nor ending the product the result is achieved by

$$Ra_0uQua_0P_1a_0uQuua_0P_2a_0uQuuua_0S$$
$$produces$$
$$Ra_0uQua_0P_2a_0uQuua_0P_1a_0uQuuua_0S.$$

Here Q must consist of u's only. For otherwise a_0uQua_0 and a_0uQuua_0 would have their initial a_0's followed by identical maximal u sequences. The u sequences uQu, $uQuu$ and $uQuuu$ are then maximal, and differ in length by one each. P_1 and P_2 again then are consecutive factors of the product. Finally, for two factors ending a product, the last production, rewritten with a_0S deleted, suffices.

The next production to be added to the bases of S_2 allows us to pass from the assertion of a product to the assertion of the first factor of a product, and hence, with the help of the previous three productions, to the assertion of an arbitrary factor of a product. The production is simply

$$ua_0Pa_0uua_0R \quad produces \quad P.$$

In translating the operations of S_1 into operations within products of S_2, we allow for passing from a product whose initial factors can be identified with the premises of an S_1 operation, to that product with the conclusion of the S_1 operation as additional factor. That additional factor must end the new product so as not to disturb the progression of the maximal u sequences. Let "$G_1, G_2, \cdots, G_k$ produce G" represent any one of the S_1 operations. Let H represent

$$ua_0G_1a_0uua_0G_2a_0\ldots\underbrace{u\ldots u}_{k}a_0G_k\underbrace{a_0u\ldots u}_{k+1}.$$

Then the corresponding S_2 operation may be represented by

$$Ha_0Ra_0uQua_0Sa_0uQuu \quad \textit{produces}$$
$$Ha_0Ra_0uQua_0Sa_0uQuua_0Ga_0uQuuu.$$

Note that the operational variables of this production are those of the S_1 production, and Q, R, S. Since each operational variable in G occurs in at least one of the G_i's, our new production will indeed have the same operational variables in its conclusion as in its premise. The portion of the premise following H insures that Q consists of u's only. This, with the form of H, insures that $G_1, G_2, \cdots, G_k$ are determined factors of the premise and G of the conclusion. Hence our transformation of the S_1 production is valid. The additional operational variables R and S require an assertion to which this production is applied to have at least $k+2$ factors, a requirement secured below. Of course, the basis of S_2 is to have the correspondent of each of the operations in the basis of S_1.

With S_1 having κ productions, the above $\kappa+4$ productions constitute all of the productions in the basis of S_2. Its sole primitive assertion is then formed as follows. Let L be the largest number of premises occurring in any production of S_1. If S_1 has λ primitive assertions, let each be repeated L times to give λL sequences each involving no other letters than $a_1, \cdots, a_\mu$. If $\lambda L < L+2$, or $\lambda L < 3$, again duplicate one of these sequences the one or two times needed to avoid these inequalities. If then $k_1, k_2, \cdots, k_M$ are these duplicated primitive assertions of S_1, the primitive assertion of S_2 will be their product

$$ua_0k_1a_0uua_0k_2a_0\ldots\underbrace{u\ldots u}_{M}\,a_0k_Ma_0\underbrace{u\ldots u}_{M+1}.$$

Now it is readily proved by induction that if at a certain point of the process for obtaining assertions in S_1 a certain finite set of assertions has been obtained, then there will be asserted in S_2 a product among whose factors are each of the above assertions repeated L times. For the primitive assertions of S_1, this is insured by the primitive assertion of S_2. Assume it to be true for the deductive process in S_1 at an arbitrary point, let P_2 be the corresponding assertion in S_2, P_1 the next assertion obtained in S_1, $P_{11}, P_{12}, \cdots, P_{1k}$ the premises of the production of S_1 yielding conclusion P_1. Then each P_{1j} appears as a factor of P_2 indeed L times at least. Hence from P_2, by the first three productions of S_2, an assertion P'_2 can be obtained in which the first k factors are $P_{11}, P_{12}, \cdots, P_{1k}$ respectively, whatever repetitions may occur among those P's. The production of S_2 corresponding to the one of S_1 in

question will then add P_1 as factor to P'_2. Mere repetition of the application of this production will then yield P''_2, which will be P'_2 with L additional factors equal to P_1. The induction is thus established. It follows that for each assertion P_1 in S_1 there will be an assertion P_2 in S_2 having P_1 as factor. By the first three productions of S_2 this factor can be made the first factor of an assertion in S_2, and hence, by the fourth production of S_2, P_1 itself will be an assertion of S_2. That is, every assertion of S_1 is an assertion of S_2. Our basis for S_2 shows that the only other assertions of S_2 are products of assertions of S_1, and so not wholly written on the letters of S_1. Hence, an enunciation of S_1 is an assertion of S_1 when and only when it is an assertion of S_2, whence the reduction of S_1 to S_2.

In our second reduction of the canonical form the productions, all with single premises by the previous reduction, now take the more special form

$$g_1P_1g_2P_2\cdot\cdot\cdot g_mP_mg_{m+1}$$

produces

$$\bar{g}_1P_1\bar{g}_2P_2\cdot\cdot\cdot\bar{g}_mP_m\bar{g}_{m+1}$$

where, however, m, and of course the g's, may vary from operation to operation. By contrast, in the previous productions P's could be repeated, have different arrangements in premise and conclusion, and in part be missing from the conclusion while present in the premise.

Again let the primitive letters of the given system be symbolized $a_1, a_2, \cdot\cdot\cdot, a_\mu$. Let its i-th production be

$$g_1P_{j_1}g_2P_{j_2}\cdot\cdot\cdot g_mP_{j_m}g_{m+1}$$

produces

$$g'_1P_{j'_1}g'_2P_{j'_2}\cdot\cdot\cdot g'_{m'}P_{j'_{m'}}g'_{m'+1}$$

where it is understood that each letter except P has i for additional subscript. The subscripts of the P's need not be distinct in premise or conclusion, while the different subscripts of the P's in the conclusion all appear in the premise. However, the letter P occurs exactly $m + m'$ times in the production.

We introduce a new primitive letter u, and for each such production two new primitive letters v_i, w_i. In obtaining the effect of the i-th production we shall, as above, leave this subscript i understood. v_i will be used in passing from an assertion involving a's only that could be the premise of the i-th production to one which has both that premise and corresponding conclusion recognizable within it; w_i in passing from such a composite assertion to the

desired conclusion only. The efficacy of our method will depend on each assertion in the new system which involves v or w having that letter only at the beginning of the assertion, and in the first case always involving exactly $2m + m'$ u's, in the second, m' u's. Our new productions will in every case explicitly exhibit this v and $2m + m'$ u's, or w and m' u's, so that we can be sure that in their application the operational variables can represent sequences of a's only. Except for a minor preliminary type, all of our "v-assertions" will be in the form

$$vug_1P_1uQ_1ug_2P_2uQ_2 \cdots ug_mP_muQ_mg_{m+1}ug'_1Q_{m+1}ug'_2Q_{m+2} \cdots ug'_{m'}Q_{m+m'}g'_{m'+1},$$

and when so asserted will have the following properties. The sequence of a's $g_1P_1Q_1g_2P_2Q_2 \cdots g_mP_mQ_mg_{m+1}$ is an assertion of the given, and indeed new system, while the sequences of a's $Q_1, Q_2, \cdots, Q_m, Q_{m+1}, \cdots, Q_{m+m'}$ can, in order, be identified with $P_{j_1}, P_{j_2}, \cdots, P_{j_m}, P_{j'_1}, \cdots, P_{j'_{m'}}$, that is, any two Q's corresponding to P's with identical subscripts are equal. Note that with all $2m + m'$ u's exhibited, the g's being given, the P's and Q's of such an assertion are uniquely identifiable in the assertion. Our method depends on the fact that when such an assertion is obtained in which the P's are null, then, due to the equalities forced on the Q's, $g_1Q_1g_2Q_2 \cdots g_mQ_mg_{m+1}$ becomes an assertion on a's only that can be identified with the premise of the i-th production of the given system, and hence $g'_1Q_{m+1}g'_2Q_{m+2} \cdots g'_{m'}Q_{m+m'}g'_{m'+1}$ an expression on a's only that will be the corresponding conclusion. Of course, each production about to be described is directly seen to be in the desired newly simplified form.

Since a null assertion has been excluded from our systems, each assertion of the given system is of the form a_jP, $j = 1, 2, \cdots \mu$. The productions

$$a_jP \quad \textit{produces} \quad va_jPu \ldots u$$

with $2m + m'$ u's in $u \ldots u$ changes each "a-assertion," i. e., assertion involving a's only, into what we shall call the intermediate v form. As all other assertions of our new system will begin with v or w, these productions will be inapplicable to them. If now an a-assertion can be the premise of the i-th production, its intermediate v form will be put into primary v form, or just v form, by the production

$$vg_1P_1g_2P_2 \cdots g_mP_mg_{m+1}u \ldots u$$

$$\textit{produces}$$

$$vug_1P_1uug_2P_2uu \ldots g_mP_mug_{m+1}ug'_1ug'_2u \ldots ug'_{m'}g'_{m'+1}.$$

Of course this production may be applicable without the P's being identifiable with those of the premise of the i-th production. But, comparing this conclusion with our general v form, we see that it satisfies the requirement thereof with all Q's null. Now any set of a-sequences that could be identified with the $P_{j_1}, P_{j_2}, \cdots, P_{j_m}, P_{j'_1}, \cdots, P_{j'_{m'}}$ of the i-th production can be built up as follows. Start with the set of null sequences. Let $Q_1, Q_2, \cdots, Q_m, Q_{m+1}, \cdots, Q_{m+m'}$ be any such derived set of a-sequences. Let $Q_{j_1}, Q_{j_2}, \cdots, Q_{j_\nu}$, j's increasing, be any subset thereof corresponding to all P's with subscripts equal to a given subscript, a_j any one of the primitive a's. Then $\cdots, a_jQ_{j_1}, \cdots, a_jQ_{j_2}, \cdots, a_jQ_{j_\nu}, \cdots$, all other Q's unchanged, will also be such a set of a-sequences. Rewrite the subscript sequence $j_1, j_2, \cdots, j_\nu$ in the form $j_1, \cdots, j_\lambda, j_{\lambda+1}, \cdots, j_\nu$ so that $j_\lambda \leqq m$, $j_{\lambda+1} > m$, and let $j_{\lambda+1} - m = j'_1, \cdots, j_\nu - m = j'_{\lambda'}$. Of course we may have $\lambda = \nu$. Now for each such choice of original P subscript, and each a_j, introduce the production

$$v \ldots ug_{j_1}P_{j_1}a_juQ_{j_1} \cdots ug_\lambda P_{j_\lambda}a_juQ_{j_\lambda} \cdots ug'_{j'_1}Q_{j_{\lambda+1}} \cdots ug'_{j'_{\lambda'}}Q_{j_\nu} \cdots$$

produces

$$v \ldots ug_{j_1}P_{j_1}ua_jQ_{j_1} \cdots ug_\lambda P_{j_\lambda}ua_jQ_{j_\lambda} \cdots ug'_{j'_1}a_jQ_{j_{\lambda+1}} \cdots ug'_{j'_{\lambda'}}a_jQ_{j_\nu} \cdots$$

all of the rest of both premise and conclusion being as in the type v form above. Such a production will then change a valid v form into a valid v form, the effect being however to "drain" the P's of such a form and "swell" the Q's. If then an assertion of the given system can be put in the form of the premise of the i-th production, the corresponding intermediate v form will pass into a v form such that successive application of the above productions will completely drain the P's thereof; and, indeed, conversely. This marks the end of the first half of the passage from a-assertion to a-assertion in the new system. While the second half could be set up by means of similar w productions in reverse, with interchange of emphasis on premise and conclusion of the i-th production, the following method is simpler. With P's all null, the v form determines the desired a-conclusion as described above. The w forms, about to be introduced, each have exactly m' u's all explicitly appearing in the productions. From such a v form with P's all null the first w form is obtained via

$$vug_1uQ_1ug_2uQ_2 \cdots ug_muQ_mg_{m+1}ug'_1Q_{m+1}ug'_2Q_{m+2} \cdots ug'_{m'}Q_{m+m'}g'_{m'+1}$$

produces

$$wg_1Q_1g_2Q_2 \cdots g_mQ_mg_{m+1}ug'_1Q_{m+1}ug'_2Q_{m+2} \cdots ug'_{m'}Q_{m+m'}g'_{m'+1}.$$

We can now get rid of the no longer interesting part of this w form, i. e., the part between w and the first u thereof, by the μ productions

$$wa_jPug'_1P_1ug'_2P_2\cdots ug'_{m'}P_{m'}g'_{m'+1}$$
$$\textit{produces}$$
$$wPug'_1P_1ug'_2P_2\cdots ug'_{m'}P_{m'}g'_{m'+1}$$

iteratively applied till letter by letter what was the original a-assertion disappears. The desired a-conclusion then would be obtained via

$$wug'_1P_1ug'_2P_2\cdots ug'_{m'}P_{m'}g'_{m'+1}$$
$$\textit{produces}$$
$$g'_1P_1g'_2P_2\cdots g'_{m'}P_{m'}g'_{m'+1}.$$

Our final system will then be on the primitive letters $a_1,\cdots,a_\mu, u, v_1, w_1, v_2, w_2,\cdots, v_\kappa, w_\kappa$, κ being the number of productions of the given system. The one primitive assertion of the new system will be the one primitive assertion of the given system, the productions of the new system, all of the above productions for each of the κ productions of the given system. Our above analysis then easily shows that the assertions of the new system involving no other letters than $a_1,\cdots,a_\mu$ are exactly the assertions of the given system, and the desired reduction has been effected.

Our third and penultimate simplifying reduction of the canonical form is to one where the operations are of the form

$$g_1Pg_2$$
$$\textit{produces}$$
$$\bar{g}_1P\bar{g}_2,$$

i. e., involve but a single operational variable. Again let a system in the previous simplified form have primitive letters $a_1, a_2,\cdots, a_\mu$, and κ operations, the number of P's in the premise, and hence conclusion, of the i-th operation being m_i. For the i-th operation, with $i=1,2,\cdots,\kappa$, and each primitive letter a_j we introduce $2m_i+1$ new primitive letters $a'_{ji}, a''_{ji},\cdots, a^{(2m_i)}_{ji}$, $a^{(2m_i+1)}_{ji}$. We also introduce the primitive letter a_{0i} and its $2m_i+1$ primed equivalents. With one such operation in mind at a time we shall, as above, omit the extra subscript i. Apart from the use of a_0 and $a^{(j)}{}_0$'s, needed to take care of g's or P's that are null, the essence of our method is to pass from an a-assertion in the form $g_1P_1g_2P_2\cdots g_mP_mg_{m+1}$ to an assertion $g'_1g'''_2\cdots g^{(2m+1)}_{m+1}P''_1P^{IV}{}_2\cdots P^{(2m)}_m$ where the superscript k say indicates that each a_j in the corresponding expression is here written $a_j{}^{(k)}$. As a result our premise will now have the form gP with $g=g'_1g'''_2\cdots g^{(2m+1)}_{m+1}$, $P=P''_1P^{IV}{}_2\cdots P^{(2m)}_m$.

In detail, we first introduce μ productions

$$a_jP \quad \textit{produces} \quad a_0a_jP, \qquad j = 1, 2, \cdots, \mu,$$

which will be applicable in fact only to assertions on $a_1, a_2, \cdots, a_\mu$, and changes any such assertion Q into a_0Q. We then introduce a finite series of finite sets of production depending in number on m and μ. The first set has the one production

$$a_0g_1P \quad \textit{produces} \quad a'_0g'_1a_0Pa''_0.$$

Inductively let the conclusion of the sole production in the $(2k-1)$-st set be in the form $G_ka_0Pa^{(2k)}{}_0$. Then the $(2k)$-th set has the μ productions

$$G_ka_0a_jP \quad \textit{produces} \quad G_ka_0Pa^{(2k)}{}_j, \qquad j = 1, 2, \cdots, \mu,$$

the $(2k+1)$-st set the sole production

$$G_ka_0g_{k+1}P \quad \textit{produces} \quad G_ka^{(2k+1)}{}_0g^{(2k+1)}{}_{k+1}a_0Pa^{(2k+1)}{}_0.$$

This is to hold for $1 \leqq k < m$, while for $k = m$ the sole production of the $(2m+1)$-st set is to be

$$G_ma_0g_{m+1}a''_0P \quad \textit{produces} \quad G_ma^{(2m+1)}{}_0g^{(2m+1)}{}_{m+1}a''_0P.$$

We then readily see that starting with an assertion on $a_1, \cdots, a_\mu$ in the form $g_1P_1g_2P_2 \cdots g_mP_mg_{m+1}$, one can, with the aid of these productions, obtain as an assertion

$$a'_0g'_1a'''_0g'''_2 \cdots a^{(2m-1)}{}_0g^{(2m-1)}{}_ma^{(2m+1)}{}_0g^{(2m+1)}{}_{m+1}a''_0P''_1a^{\mathrm{IV}}{}_0P^{\mathrm{IV}}{}_2 \cdots a^{(2m)}{}_0P^{(2m)}{}_m.$$

Furthermore, note that starting with an assertion on $a_1, a_2, \cdots, a_\mu$, flanked on the left by a_0 as above, one can apply the above operations only in the following order, if at all. First, the sole operation of the first set; and inductively, if the operation in the $(2k-1)$-st set has last been applied, the next applicable operation can only be an operation in the $2k$-th set or the operation in the $(2k+1)$-st set, if an operation in the $(2k)$-th set has last been applied, the next applicable operation can only be an operation in the same set, or the operation in the next set. Furthermore, the last operation in its premise explicitly indicates the a''_0, first introduced into an assertion only as a result of the first operation. It readily follows that if the last operation does enter into a possible sequence of operations, the conclusion thereof can have no letter a_j in it without a superscript. The entire given assertion has thus been translated; and it is readily seen that that last assertion, and hence given assertion, are and can be put in the forms above given.

The actual correspondent of the original i-th operation in translated form may then be written simply

$$a'_0g'_1a'''_0g'''_2\cdot\cdot\cdot a^{(2m-1)}{}_0g^{(2m-1)}{}_m a^{(2m+1)}{}_0g^{(2m+1)}{}_{m+1}P$$

produces

$$a'_0\bar{g}'_1a'''_0\bar{g}'''_2\cdot\cdot\cdot a^{(2m-1)}{}_0\bar{g}^{(2m-1)}{}_m a^{(2m+1)}{}_0\bar{g}^{(2m+1)}{}_{m+1}P;$$

and the passage from this translated conclusion to the actual conclusion can be effected by a set of productions the reverse of those above given. That is, in each of the above productions prior to the actual correspondent of the i-th production replace all g_j's by $\bar{g}_j$'s, and *interchange hypothesis and conclusion.* The resulting productions then clearly suffice to yield the conclusion yielded by the original i-th production. True, the complete set of productions thus set up to take the place of the original i-th production may now allow other paths than from assertion on $a_1,\cdot\cdot\cdot,a_\mu$, down the first group of productions, through the intermediate production, and up the second group of productions to new assertion on $a_1,\cdot\cdot\cdot,a_\mu$.[17] But it is readily seen that any departures from this progression merely constitute unravelings of parts of such a progression, or, apart from such unravelings, constitute shortcuts of valid full progressions of this type. Since, furthermore, one can change the set of productions one is working with only when an assertion on $a_1,\cdot\cdot\cdot,a_\mu$ alone is obtained, the validity of our reduction follows.

Our final reduction is to a system whose operations are in the form

$$gP$$
$$\textit{produces}$$
$$Pg'.$$

The present method assumes that in the productions of the previous system, all in the form

$$g_1Pg_2$$
$$\textit{produces}$$
$$g'_1Pg'_2,$$

g_1 and g_2 are never null. We therefore actually first need the following preliminary reduction. Introduce a new primitive letter a_0, and if h is the sole primitive assertion of the given system let a_0ha_0 be the sole primitive assertion of the new system. Replace each of the above operations of the given system by

$$a_0g_1Pg_2a_0 \quad \textit{produces} \quad a_0g'_1Pg'_2a_0$$

and finally add the production

[17] This could have been avoided say by the v, w method used earlier.

a_0Pa_0 *produces* P.

Except for the last production the new system may be said to be simply isomorphic with the old, P being an assertion in the given system when and only when a_0Pa_0 is an assertion in the new system. The last operation then merely recovers the assertions of the given system. Note that even that last operation is in the desired form with neither g_1 nor g_2 null.

Assume then that such is our given system with primitive letters again $a_1, a_2, \cdots, a_\mu$. We introduce new primitive letters $\bar{a}_1, \bar{a}_2, \cdots, \bar{a}_\mu$, and "translating productions"

$$a_jP \text{ produces } P\bar{a}_j, \quad \bar{a}_jP \text{ produces } Pa_j, \qquad j = 1, 2, \cdots, \mu.$$

Starting with an assertion of the form $a_{i_1} \cdots a_{i_j} a_{i_{j+1}} \cdots a_{i_n}$, these productions will yield only assertions of the form $a_{i_{j+1}} \cdots a_{i_n}\bar{a}_{i_1} \cdots \bar{a}_{i_j}$, $\bar{a}_{i_1} \cdots \bar{a}_{i_j}\bar{a}_{i_{j+1}} \cdots \bar{a}_{i_n}$, $\bar{a}_{i_{j+1}} \cdots \bar{a}_{i_n}a_{i_1} \cdots a_{i_j}$, in addition to the original assertion. Only one of these $2n$ distinct forms consists wholly of unbarred letters, i. e., the original form, while continued application of the above operations merely keeps deriving these $2n$ "equivalent forms" cyclically, so that anyone can thus be obtained from any other.

Our reduction will then be effected if for each operation "g_1Pg_2 *produces* $g'_1Pg'_2$" of the given system we introduce in the new system the operation

$$\bar{g}_2g_1P$$
$$\textit{produces}$$
$$Pg'_2\bar{g}'_1,$$

where $\bar{g}_2$, for example, is g_2 with each letter replaced by the corresponding barred letter. Of course, the one primitive assertion of the given system is also the one primitive assertion of the new system. Note that if at any point an assertion without barred letters appears, then if it can be written g_1Pg_2, the first given translating operations derive from it $\bar{g}_2g_1P$, hence the above yields $Pg'_2\bar{g}'_1$, and so finally $g'_1Pg'_2$ is obtained as desired. That is, the new system contains all of the assertions of the given system. It further follows that the assertions of the new system consist only of the assertions of the given system and their equivalents. For, proceeding inductively, this clearly remains true under the translating operations. Now suppose it is true of an assertion in the form $\bar{g}_2g_1P$. Since $\bar{g}_2$ and g_1 are not null, $\bar{g}_2g_1$ alone exhibits a change from barred to unbarred letters. P therefore must consist of unbarred letters only. $\bar{g}_2g_1P$ is therefore a translation of the assertion g_1Pg_2 of the original system, and hence the conclusion $Pg'_2\bar{g}'_1$ is a translation of $g'_1Pg'_2$, also an assertion of the original system. The desired reduction has thus been effected.

The original system in canonical form has thus been reduced to a system in normal form. At each stage in that reduction the primitive letters of the new system are the primitive letters of the preceding system and a finite number of additional letters, while the assertions of that preceding system are exactly those assertions of the new system which involve only primitive letters of the preceding system. The same is then true of the original system in canonical form and the final system in normal form, whence the theorem of the introduction.[18]

THE CITY COLLEGE,
COLLEGE OF THE CITY OF NEW YORK.

---

[18] While the present paper is presented as a contribution to the literature as it now exists, its main development, that of Section **2**, as well as the further transformation mentioned in the introduction, was obtained by the writer in substantially the form here given in the summer of 1921. The larger development of which this was a part is no longer essentially new, but may be worth a brief résumé. In the fall of 1920, starting with the formulation referred to in footnote 2 as canonical form *A*, the operation of substitution of that formulation was weakened in two successive stages to yield canonical forms *B* and *C*, the last only allowing any 1-1 replacement of variables by variables. It was first shown in detail that each of these forms was reducible to the other two, and then the project that led to the above changes of canonical form *A* was carried through, namely the reduction of what would now be termed the restricted functional calculus of *Principia Mathematica* to a particular system in canonical form *C*. Much of this work has since been found to be seriously in error, but easily corrected by the methods then employed. As a result of the last reduction it appeared obvious to the writer that all of *Principia Mathematica* could likewise be reduced to a system in canonical form *C*. In the summer of 1921, the intervening work on the problem of tag suggested the reduction of canonical form *C* to the canonical form of the present paper, and this reduction was followed by the successive reductions to normal form essentially as given in Section **2**. The added methods thus revealed led us to conclude that not only *Principia Mathematica*, but any symbolic logic whose operations could effectively be reproduced in *Principia Mathematica*, and hence probably any (finitary) symbolic logic could be reduced to a system in canonical form, and consequently to a system in normal form. But now the entire direction of our thought, that of solving the decision problem for arbitrary systems, was reversed. Having noted the identity of canonical sets and normal sets referred to in the introduction, our last conclusion was transformed into the generalization that every generated set of sequences on a finite sets of letters was a normal set. The seeming counter example furnished by the diagonal method then led to an informal proof that the decision problem for the class of systems in normal form was unsolvable. In the early fall of 1921, the formal proof of this unsolvability, referred to in the introduction, was outlined, and led to the further conclusion that not only was every (finitary) symbolic logic incomplete relative to a certain fixed class of propositions (those stating that a given sequence was or was not an assertion in a given normal system) but that every such logic was extendable relative to that class of propositions. Since the earlier formal work made it seem obvious that the actual details of the outline could be supplied, the further efforts of the writer were directed towards establishing the universal validity of the basic identification of generated set with normal set.

A CONVERGENCE PROOF INVOLVING AN INSEPARABLE MULTIPLE CONTOUR INTEGRAL.*

By CHESTER C. CAMP.

In a previous paper [1] the author considered the expansion of $f(x_1, x_2, \cdots x_p)$ in terms of solutions of the system (3), [numbers less than (22) will refer to items in the former paper [1]],

$$(3) \qquad X'_j + [\sum_{i=1}^{p} \lambda_i a_{ji}(x_j)]X_j = 0, \qquad (j = 1, 2, \cdots, p);$$

and the auxiliary conditions

$$(4) \quad X_j(a_j) + C_{j2}X_j(a_{2j}) + \cdots + C_{j,k_j-1}X_j(a_{k_j-1,j}) + C_{j,k_j}X_j(b_j) = 0,$$

where $C_{j,k_j} \neq 0$, $(j = 1, 2, \cdots, p)$. The proof of convergence was accomplished by the extension of the contour integral method.[2] But in order to make the integrals separate in pairs after appropriate transformations on the parameters it was assumed that the average value of each $a_{jk}(x_j)$ was maintained over every subinterval for the variable x_j; i. e.,

$$(11) \qquad \int_{a_{ij}}^{a_{i+1,j}} a_{jk}(x_j)\,dx_j/(a_{i+1,j} - a_{ij}) = A_{jk}, \qquad (i = 1, 2, \cdots, k_j - 1).$$

The purpose of this paper is to indicate how this restriction may be removed. The work is somewhat simplified by taking first the case $p = 2$. The changes necessary for $p > 2$ will then be obvious.

There are two things to be proved: first, that under the new conditions the residue at a simple set of principal parameter values of the function (19) will give the corresponding term of the series (15); and secondly, that the

* Presented to the Society, September, 1941; Received September 11, 1941; Revised February 10, 1942.

[1] Camp, "On multiparameter expansions associated with a differential system and auxiliary conditions at several points in each variable," *American Journal of Mathematics*, vol. 60 (1938), pp. 447-452.

[2] Camp, "Multiple Fourier series," *Transactions of the American Mathematical Society*, vol. 25 (1923), pp. 131-132.

limit of a certain multiple integral over properly oriented contours in the various λ_i-planes yields the value of f for a point at which it is continuous.

One may use the transformation of parameters given in (10) or, by making certain modifications, one may keep the original λ_i. The latter plan seems simpler since the contour integrals in ν_i would no longer be separable anyway. The necessary changes will be indicated.

To prove the first of the statements above we proceed as follows. Although the more general interpretation of a double integral in two complex variables involves a two-dimensional surface in a four-dimensional space,[3] it is found sufficient here to use the two-plane [4] representation of the independent variables λ_1, λ_2. We can then write for the residue of $P(\lambda_1, \lambda_2)/W_1W_2$ at a simple place $(\lambda^*_1, \lambda^*_2)$ at which $W_1(\lambda_1, \lambda_2) = 0$, $W_2(\lambda_1, \lambda_2) = 0$ the form

$$(22) \qquad R_1 = \left.\frac{P(\lambda_1, \lambda_2)}{\dfrac{D(W_1, W_2)}{D(\lambda_1, \lambda_2)}}\right] \quad \lambda_k = \lambda^*_k, \quad (k = 1, 2),$$

where P, W_1, W_2 are analytic functions.[5] For a place $\lambda_1^{(\rho)}, \lambda_2^{(\rho)}$ at which W_1 has a double point but $W_2 \neq 0$, one has the residue

$$(23) \qquad R_2 = \left.\frac{P(\lambda_1, \lambda_2)}{W_2(\lambda_1, \lambda_2)\left[\left(\dfrac{\partial^2 W_1}{\partial\lambda_1\partial\lambda_2}\right)^2 - \dfrac{\partial^2 W_1}{\partial\lambda_1^2}\,\dfrac{\partial^2 W_2}{\partial\lambda_2^2}\right]^{\frac{1}{2}}}\right] \quad \lambda_k = \lambda_k^{(\rho)}, \quad (k = 1, 2).$$

It is not difficult to see that the functions W_j, which will now be simply the left hand members of (4), $(j = 1, 2)$, will have in general isolated common zero-places since they are functions which are uniform, analytic, and devoid of essential singularities in the finite field of variation.[6] We assume that they are independent and free of possible vanishing common factors.

The Green's system (14) will be modified by omitting the argument ν_j in $W_j(\nu_j)$, and also the determinant $|A_{jk}|$. One may then show by proper orientation of the surface, or curves of integration in the two-plane method, and

---

[3] Cf. A. R. Forsyth, *Lectures introductory to the theory of functions of two complex variables*, 1914, Chapter VI; also Stefan Bergman, "On the surface integrals of functions of two complex variables," *American Journal of Mathematics*, vol. 63, 1941, pp. 295-318.

[4] Cf. Forsyth, *loc. cit.*, p. 161, sections 99, 100.

[5] Cf. H. Poincaré, "Sur les résidus des intégrales doubles," *Acta Mathematica*, vol. 9, 1887, p. 351, where these results are given for polynomial functions. See also Bergman, *loc. cit.*, p. 315.

[6] Cf. Forsyth, *loc. cit.*, p. 209.

the use of (22) for the ordinary case of simple characteristic values that the residues of (19) in the new form will give the several terms of the expansion (15) sought for $f(x_1, x_2)$. This is easily done by expressing Q of (17) in terms of the K_{ij} of (9). Q will take the form $\frac{D(W_1, W_2)}{D(\lambda_1, \lambda_2)}$ in (22) instead of the previous form (20).

One may show that the real parts of λ^*_k and $\lambda_k^{(\rho)}$ are bounded and that within finite rectangles in the λ_k-planes there will be a finite number of values of λ^*_k, $\lambda_k^{(\rho)}$. Moreover if the λ_k are uniformly bounded from the characteristic values then W_1, W_2 will be uniformly bounded away from zero.

To show that the real parts of λ^*_k are bounded consider a method similar to that used by C. E. Wilder.[7] If one defines

$$\mu_k = -\lambda_k A_{ik}(a_{2i}), \qquad (k = 1, 2), \tag{24}$$

where $i = 1$ or 2 is so chosen that $A_{ik}(a_{2i})$ is numerically the smaller non-vanishing value for the particular k, then the product $W_1 W_2$ will take the form

$$1 + Q_{o1} \exp(\mu_1 + a\mu_2) + Q_{1o} \exp(b\mu_1 + \mu_2) + \cdots + Q_{rs} \exp(c\mu_1 + d\mu_2). \tag{25}$$

Since there are a finite number of exponents there will be a dominant one for $R(\mu_1) \geqq X_1$ sufficiently large and μ_2 fixed; similarly for $R(\mu_2) \geqq X_2$ and μ_1 fixed.

If we take $R(\mu_1) = \mu R(\mu_2)$, then among the finite number of exponents in (25) there will be one with an algebraically largest (and smallest) real part which will dominate for $R(\mu_2) \geqq X \geqq X_i$, $(i = 1, 2)$, for a suitable value of μ (also for $R(\mu_2) \leqq -X$), such that the product $W_1 W_2$ divided by one of the corresponding exponentials will differ from its corresponding coefficient Q_{hk} by less than ε in absolute value. Evidently then $W_1 W_2$ cannot vanish for $|R(\mu_1)| > \mu X$, $|R(\mu_2)| > X$. Likewise one sees that λ^*_k and $\lambda_k^{(\rho)}$ have bounded real parts.

In order to finish the convergence proof it remains to prove the following lemmas:

LEMMA I.

$$\lim_{|\lambda_k| \to \infty} \frac{1}{-4\pi^2} \int_g \int d\lambda_1 d\lambda_2 / W_1 W_2 \lambda_1 \lambda_2 = 1/4.$$

[7] Cf. C. E. Wilder, "Problems in the theory of ordinary differential equations with auxiliary conditions at more than two points," *Transactions of the American Mathematical Society*, vol. 18 (1917), p. 420.

LEMMA II.

$$\lim_{|\lambda_k|\to\infty} \frac{1}{-4\pi^2}\int_S\int |\exp(-h_1\lambda_1 - h_2\lambda_2)|\, d\lambda_1 d\lambda_2 / W_1 W_2 \lambda_1\lambda_2 = 0,$$

where $0 < h_j < (b_j - a_j)(A_{j1} + A_{j2})$ *for at least one value of* j.

Here S is a dicylinder $|\lambda_k| = C_k$, $(k = 1, 2)$, and the C_k are radii of circles in the λ_k-planes, respectively, which are uniformly bounded away from the characteristic values. These circles will not usually be uniformly increasing as the contours expand. In the rare case in which W_1 or W_2 has factors of the form $1 - \exp \alpha_k\lambda_k$ and $1 - \exp \beta_k\lambda_k$ where α_k and β_k are incommensurable as well as in the ordinary case the corresponding expansion series must be arranged to conform with the sequence of expanding contours.

The proof of these lemmas may be considered by cases. Since the characteristic values have numerically bounded real parts one may replace, when convenient, the circular contours in either plane by rectangular ones on which the integrands will be bounded. For the case in which at least one of the $R(\lambda_k)$, $(k = 1, 2)$, is negative it is easy to see by former methods of proof by using at least one semi-circular arc and a convenient parabola [8] and taking absolute values that the double limit approached is zero. In Lemma II when both real parts are positive one transforms the integrand by multiplying both numerator and denominator by a suitable exponential and proceeds as before. For Lemma I both real parts will be positive for half of each circular contour. If the integrand is decomposed into the form $1/\lambda_1\lambda_2 + (1 - W_1W_2)/\lambda_1\lambda_2 W_1 W_2$ the first term integrated over this semi-circle in each λ-plane will yield $1/4$ while the second term for these parts of the dicylinder will contribute zero to the double limit. One may therefore state the following

THEOREM: [9] *Let* $f(x_1, x_2)$ *consist of a finite number of pieces, each real and possessing a continuous partial derviative in each argument for its particular subregion of the region* T: $a_j \leqq x_j \leqq b_j$, $(j = 1, 2)$; *let each* $a_{ji}(x_j)$ *be integrable and either positive or identically zero in* T; *let the determinant of the integrals of* $a_{ji}(x_j)$ *over their respective intervals be different from zero. Then the expansion analogous to the one formerly obtained* [9] *will converge at any interior point of* T *to the so-called mean value of* f. *If the terms are grouped appropriately the series will converge uniformly to* f *at any interior point at which* f *is continuous. In the rare case in which the characteristic*

[8] Cf. M. G. Carman, "A convergence proof for simple and multiple Fourier series," *Bulletin of the American Mathematical Society*, vol. 30 (1924), p. 413.

[9] Cf. Camp, *American Journal of Mathematics, loc. cit.*, p. 452, Theorem II.

places are not simple but W_1 has a double point where $W_2 \neq 0$ the usual expansion for f must be augmented by a term of the form

$$\int_{a_1}^{b_1}\int_{a_2}^{b_2} f(s_1, s_2)\,|\,a_{jt}(s_j)\,|\,R_2\,ds_2 ds_1$$

where R_2 is the residue of the revised Green's system at $\lambda_k^{(\rho)}$ formed from (23) by proper substitutions. An analogous term is added also if W_2 has a double zero at a place where W_1 does not vanish.

It is clear that this theorem can be extended immediately to the case $p > 2$, also that if the $a_{jt}(x_j)$ are allowed to change sign [10] or otherwise vanish the corresponding expansion would converge but not usually to the mean value of f.

UNIVERSITY OF NEBRASKA,
LINCOLN, NEBRASKA.

[10] Cf. Camp, "Expansion involving differential equations in which the coefficient of a parameter changes sign," *National Mathematics Magazine*, vol. 12, 1938, pp. 216-222.

A CHARACTERIZATION OF POLYNOMIAL RINGS BY MEANS OF ORDER RELATIONS.*

By HOWARD LEVI.

One of the most striking differences between the theory of polynomial rings and that of the general ring is that in the former theory algorithms of elimination and reduction are used which are not available in the latter. The contrast between Macaulay's *Algebraic Theory of Modular Systems* and van der Waerden's *Moderne Algebra* (particularly chapter XII) illustrates this point. This difference seems to be based on the possibility of introducing order relations more or less explicitly into polynomial rings.[1] In this paper we consider a ring for whose elements certain order relations and algorithms have been assumed. The assumptions made were borrowed from polynomial rings, where they appear as fairly obvious facts. What we have accomplished is to show that these facts characterize polynomial rings; that any ring for which they are postulated must be a polynomial ring. Recognizing that polynomial rings with different types of coefficient domains differ significantly, we have presented two lists of assumptions. The first leads to a ring of polynomials whose coefficients constitute a domain of integrity which contains a unit element, and in which the Hilbert basis theorem holds. The coefficient domain derived with the second list is a field. These two lists of asumptions differ only slightly, and minor modifications of them could be made which would lead to more general types of coefficients. We have not discussed these cases, believing that a polynomial whose coefficients are of a very general character has few claims to attention which it does not share with any element of a general ring.

There is a close connection between valuation theory and the contents of this paper. The rank function employed below could be viewed as a special case of the general valuation function employed by W. Krull in "Allgemeine Bewertungstheorie," *Jour. f. d. reine und ang. Math.*, vol. 187 (1932), pp. 160-191, and others. The principal difference between the present exposition and that of valuation theory is that in the latter the values of the valuation

* Received December 5, 1941; Presented to the American Mathematical Society, October 25, 1941.

[1] A very explicit use of order relations can be found in F. S. Macaulay, "Some properties of enumeration in the theory of modular systems," *Proceedings of the London Mathematical Society*, vol. 26 (1927), pp. 531-555. A considerably less explicit ordering is found in Macaulay, *Algebraic Theory of Modular Systems*, p. 7.

function are elements of an ordered group, whereas in our case the values of the rank function are elements of an ordered set for whose elements no composition is defined or required.

The Assumptions.

1. We consider a *commutative ring* $\mathfrak{R}$, a *fully ordered set* $\mathfrak{M}$, and a function μ which maps $\mathfrak{R}$ onto $\mathfrak{M}$. We assume that $\mathfrak{R}$ contains more than one element. Capital italic letters are used exclusively to denote elements of $\mathfrak{R}$. Small italic letters denote positive or negative integers except when used as exponents; they then denote positive integers. The quantity $\mu(A)$ is referred to as the *mark* of A. If $\mu(A) < \mu(B)$ we say that A is *lower* than B, and B is *higher* than A. Of any two elements with distinct marks, one must be lower than the other. These relations are transitive.

2. We make the following assumptions connecting $\mathfrak{R}$, μ, $\mathfrak{M}$. They are understood to hold for all elements of $\mathfrak{R}$ not explicitly excluded in their formulation.

I(a). If $A \neq 0$, then $\mu(0) < \mu(A)$.

I(b). If $A \neq 0$, then $\mu(B) \leq \mu(AB)$.

I(c). If $A \neq 0$, and $\mu(B) < \mu(C)$, then $\mu(AB) < \mu(AC)$.

I(d). If $A \neq 0$, and $\mu(B) = \mu(C)$, then $\mu(AB) = \mu(AC)$.

I(e). If $\mu(A + B) < \mu(A)$, then $\mu(A) = \mu(B)$.

II. If, for some A, B, m, $\mu(A) \leqq \mu(B) < \mu(A^m)$, then there exist elements P, Q, R such that

$$\mu(A + PR) < \mu(A), \qquad \mu(B + QR) < \mu(B), \qquad \mu(P) < \mu(A).$$

III. For every infinite sequence of non-zero elements $A_1, A_2, \cdots$, there exists an integer m, and elements $C_1, C_2, \cdots, C_m$, such that

$$\mu(A_{m+1} + C_1A_1 + C_2A_2 + \cdots + C_mA_m) < \mu(A_{m+1}),$$

and

$$\mu(C_iA_i) \leq \mu(A_{m+1}), \qquad i = 1, \cdots, m.$$

We shall be occupied principally with these assumptions. Later, we shall consider briefly another set, composed of I(a) — I(e), together with

II′. If, for some A, B, m, $\mu(A) < \mu(B) < \mu(A^m)$, then there are elements P, Q, R such that

$$\mu(A) = \mu(PR), \qquad \mu(B) = \mu(QR), \qquad \mu(P) < \mu(A).$$

III′. For every infinite sequence $A_1, A_2, \cdots$ of non-zero elements there exist positive integers i and j, and an element C such that $\mu(CA_i) = \mu(A_{i+j})$.

IV′. If $A \neq 0$ and $\mu(A) = \mu(B)$ there is a C such that $\mu(A + CB) < \mu(A)$.

Using the first list of assumptions, we show that $\mathfrak{R}$ contains a subring $\mathfrak{R}_0$ and elements $X_1, \cdots, X_n$, algebraically independent over $\mathfrak{R}_0$, and such that $\mathfrak{R}$ is the polynomial ring $\mathfrak{R}_0[X_1, \cdots, X_n]$. $\mathfrak{R}_0$ is shown to be a domain of integrity in which the Hilbert basis theorem holds. It is also shown that $\mathfrak{R}_0$ contains a unit element. The assumptions of the second list are stronger than and imply those of the first. They yield the further result that $\mathfrak{R}_0$ is a field. As for the ordered set $\mathfrak{M}$, it is shown to consist of an element $\mu(0)$, an element $\mu(1)$ which is the mark of all non-zero elements of $\mathfrak{R}_0$, and of the totality of marks of all power products of $X_1, \cdots, X_n$. The ordering is shown to be such that $\mu(0)$ is less than $\mu(1)$, which is less than the mark of any power product of the X_i. The power products are compared in the usual dictionary manner. Those elements of $\mathfrak{R}$ whose marks are not already accounted for are polynomials in which the X_i figure effectively. The mark of such a polynomial is the mark of the highest power product present in it effectively.

3. It is a relatively simple matter to verify that if $\mathfrak{R}_0$ is any domain of integrity which contains a unit and in which the Hilbert basis theorem holds, and if $X_1, \cdots, X_n$ are unknowns, then $\mathfrak{R}_0[X_1, \cdots, X_n]$, when ordered as above, satisfies the assumptions of the first list. Certainly assumptions I(a)-I(e) are satisfied by such a ring. Verification of assumption II is immediate, once the inequalities are restated in terms of exponents. Assumption III is somewhat more complex. One needs the fact that every infinite set of power products of finitely many letters contains a sequence each of whose terms divides all its successors.[2] This fact, used together with the above conventions concerning the ordering, and with the fact that in $\mathfrak{R}_0$ the Hilbert basis theorem holds, makes it easy to show that such a ring satisfies assumption III as well. It is also readily seen that if $\mathfrak{R}_0$ is a field, the ring satisfies the assumptions of the second list. We shall not discuss this subject further. It is with the converse of these statements that this paper is concerned.

4. We now establish some immediate consequences of our assumptions.

(1) *$\mathfrak{R}$ is a domain of integrity.* This follows from I(a) and I(b).

[2] When stated in terms of polynomials, the first part of assumption III bears a certain resemblance to the Hilbert basis theorem. The second part, however, is not

(2) $\mu(A) = \mu(-A)$. This is obvious if $A = 0$. If $A \neq 0$ write $0 = A + (-A)$. By I(a), $\mu(0) < \mu(A)$, whence, by I(e), $\mu(A) = \mu(-A)$.

(3) *If* $\mu(A+B) > \mu(A)$, *then* $\mu(A+B) = \mu(B)$. Write $A = (A+B) + (-B)$. By I(e), the relation $\mu(A) < \mu(A+B)$ implies $\mu(A+B) = \mu(-B)$. By (2), then, $\mu(A+B) = \mu(B)$.

(4) $\mu(A+B+\cdots+F) \leqq \max(\mu(A), \mu(B), \cdots, \mu(F))$. This follows from (3) for the case of two summands. It follows by induction for any larger number of summands, for then $\mu(A+B+\cdots+F) \leqq \max(\mu(A), \mu(B+\cdots+F))$. We have, as a special case, $\mu(nA) \leqq \mu(A)$, $n = 1, 2, \cdots$.

(5) $\mu(A) = \mu(B)$ *and* $\mu(C) = \mu(D)$ *imply* $\mu(AC) = \mu(BD)$. If one of the four letters is zero this is obvious. If none of them is zero then I(d) applies. Used twice, it yields $\mu(AC) = \mu(BC)$ and $\mu(BC) = \mu(BD)$, which implies the result.

(6) $\mu(AB) = \mu(AC)$ *and* $A \neq 0$ *imply* $\mu(B) = \mu(C)$. The other possibilities are excluded by I(c).

(7) $\mathfrak{M}$ *is well-ordered.* If this were not so there would exist an infinite sequence $A_1, A_2, \cdots$, with $\mu(A_1) > \mu(A_2) > \cdots$. Invoking assumption III, we have an integer m and elements $C_1, \cdots, C_m$ for which

$$\mu(A_{m+1} + C_1A_1 + \cdots + C_mA_m) < \mu(A_{m+1}) \text{ and } \mu(C_iA_i) \leqq \mu(A_{m+1}), \quad i = 1, 2, \cdots, m.$$

Obviously the first inequality could not hold if all the C_i were zero. But if one of the C_i were different from zero we should have, by I(b), $\mu(C_iA_i) \geqq \mu(A_i)$, whence $\mu(C_iA_i) > \mu(A_{m+1})$. This contradiction proves the statement.

The Ring of Coefficients.

5. We now single out a subset of $\mathfrak{R}$, the set of *coefficients*. *Zero* is a coefficient. A non-zero element X of $\mathfrak{R}$ is a coefficient *if and only if* $\mathfrak{R}$ *contains a non-zero element* A *such that* $\mu(XA) = \mu(A)$. We denote the totality of coefficients by $\mathfrak{R}_0$.

Theorem. $\mathfrak{R}_0$ *is a subring of* $\mathfrak{R}$. *It is a domain of integrity with unit element in which the Hilbert basis theorem holds. All non-zero elements of* $\mathfrak{R}_0$ *have the same mark, which mark is less than the mark of any element of* $\mathfrak{R}$ *not a coefficient. Each non-zero coefficient* X *has the property that for all* Y, $\mu(Y) = \mu(XY)$.

---

contained in that theorem. Actually assumption III is a generalization of assumption III′, and III′ is entirely different from the Hilbert basis theorem.

First we show that if X is a coefficient and Y is such that $\mu(Y) < \mu(X)$, then Y is zero. We have, for some non-zero A, $\mu(XA) = \mu(A)$. Using I(c), the hypothesis $A \neq 0$ implies $\mu(AY) < \mu(AX)$, whence $\mu(AY) < \mu(A)$. If Y were not zero, this would contradict I(b). Thus all non-zero coefficients have the same mark, and this mark is not greater than the mark of any element not a coefficient. Let this mark be denoted by $\mu(1)$.

We show next that any element whose mark is $\mu(1)$ is a coefficient, so that the mark of any element not a coefficient exceeds $\mu(1)$. Let X be a non-zero coefficient, and let Y be an element whose mark is that of X. There is a non-zero A such that $\mu(A) = \mu(XA)$. But from $\mu(X) = \mu(Y)$ we have $\mu(AX) = \mu(AY)$. Thus $\mu(A) = \mu(AY)$ and Y is a coefficient.

6. We show further that non-zero coefficients exist, and that $\mathfrak{R}$ contains a unit element (itself a non-zero coefficient). Consider that subset of $\mathfrak{M}$ obtained by deleting $\mu(0)$. Since $\mathfrak{M}$ is well-ordered, this set has a least mark, and there is an element A which possesses this mark. Consider the infinite sequence $A, A, \cdots$, to which assumption III may be applied. We have $\mu(A + CA) < \mu(A)$, where CA is the sum of finitely many products C_iA. In virtue of assumption I(e) we have $\mu(A) = \mu(CA)$, whence, since A is not zero, C is not zero. This shows that C is a non-zero coefficient. The last equation also shows that $\mu(C) \leq \mu(A)$ (in view of I(b)); it follows then from the minimal character of A that $\mu(C) = \mu(A)$. Multiplying by C the terms in the relation $\mu(A + CA) < \mu(A)$ and applying I(c), we have $\mu((C^2 + C)A) < \mu(AC)$. It follows that $\mu(C^2 + C) < \mu(C)$, and therefore $C^2 + C = 0$. This equation shows that $(-C)$ is the required unit element.

7. It can now be established that the elements of $\mathfrak{R}_0$ constitute a ring. Let X and Y denote non-zero coefficients, and let A and B be non-zero elements such that $\mu(A) = \mu(XA)$, $\mu(B) = \mu(YB)$. Then, by (5), $\mu(AB) = \mu(ABXY)$. Since, by (1), AB is not zero, XY is a coefficient. It is easy to see that $X + Y$ is a coefficient for by (4), $\mu(X + Y) \leq \max(\mu(X), \mu(Y)) = \mu(1)$. That $-X$ is a coefficient follows from (2). This completes the proof that $\mathfrak{R}_0$ is a ring. $\mathfrak{R}_0$ is certainly a domain of integrity, because it is a subring of the domain of integrity $\mathfrak{R}$.

8. We complete the proof of the theorem by showing that the Hilbert basis theorem holds in $\mathfrak{R}_0$. We prove its equivalent, the ascending chain condition. Let $\mathfrak{A}_1, \mathfrak{A}_2, \cdots$ be an infinite sequence of ideals in $\mathfrak{R}_0$, each properly contained in its successors. Let A_1 be any element of $\mathfrak{A}_1$, and, for each subscript k, let A_{k+1} be an element of $\mathfrak{A}_{k+1}$ not contained in $\mathfrak{A}_k$. By assumption III, there is an integer m and elements $C_1, \cdots, C_m$ such that

$$\mu(A_{m+1} + C_1A_1 + \cdots + C_mA_m) < \mu(A_{m+1}).$$

Since $\mu(A_{m+1}) = \mu(1)$, which does not exceed the mark of any non-zero element of R, we must have $A_{m+1} + C_1A_1 + \cdots + C_mA_m = 0$. This equation would contradict the specification that A_{m+1} does not belong to $\mathfrak{J}_m$, if it were known that the C_i belonged to $\mathfrak{R}_0$. We show that all the C_i must belong to the $\mathfrak{R}_0$, with the possible exception of C_1, and that if A_1 is not zero, C_1 must likewise belong to $\mathfrak{R}_0$. This is sufficient to complete the proof. Observe that the choice of the A_i was such that only A_1 could be zero. Of course III does not assert that the C_i are in $\mathfrak{R}_0$. However, it does assert that $\mu(C_iA_i) \leq \mu(A_{m+1}) = \mu(1), i = 1, \cdots, m$. These relations imply that $\mu(C_i) \leq \mu(1)$ for each i for which A_i is not zero.

The Set of Marks.

9. We turn now to the set $\mathfrak{M}$ and investigate those of its members which are different from $\mu(0)$ and $\mu(1)$. If there are no such members, then $\mathfrak{R}$ and $\mathfrak{R}_0$ coincide, and while this eventuality is consistent with our assumptions, it renders them trivial. We assume henceforth that $\mathfrak{R}$ contains elements which are not coefficients.

10. Let $\mathfrak{S}_0$ denote the totality of such elements and let $\mu(\mathfrak{S}_0)$ denote the totality of their marks. We have $\mathfrak{S}_0 = \mathfrak{R} - \mathfrak{R}_0$. Because $\mathfrak{M}$ is well-ordered, $\mu(\mathfrak{S}_0)$ contains a least mark, so that $\mathfrak{S}_0$ contains an element A_1 which is not higher than any other element of $\mathfrak{S}_0$. Let $\mathfrak{R}_1$ be the totality of elements X of $\mathfrak{S}_0$ for which there exist elements A of $\mathfrak{R}$ such that $\mu(X) = \mu(AA_1)$. $\mathfrak{R}_1$ is not empty, for it contains A_1 (any non-zero coefficient can serve as the A), together with all multiples of A_1 by any non-zero element of $\mathfrak{R}$. It may contain other elements. It may be that $\mathfrak{R}_1$ coincides with $\mathfrak{S}_0$. The following steps, superfluous under these circumstances, are directed at the case in which $\mathfrak{S}_0$ is not filled out by $\mathfrak{R}_1$.

11. Let $\mathfrak{S}_1$ denote the totality of elements of $\mathfrak{S}_0$ not in $\mathfrak{R}_1$. We have $\mathfrak{S}_1 = \mathfrak{S}_0 - \mathfrak{R}_1$. $\mathfrak{S}_1$ contains a lowest element. Let A_2 be such an element. Because A_2 is in $\mathfrak{S}_0$ we have $\mu(A_2) \geq \mu(A_1)$. Furthermore, the fact that A_2 is not in $\mathfrak{R}_1$ implies that the mark of A_2 is different from that of A_1. Thus $\mu(A_1) < \mu(A_2)$. Let $\mathfrak{R}_2$ be the totality of elements X of $\mathfrak{S}_1$ for which there is an A such that $\mu(X) = \mu(AA_2)$. If $\mathfrak{R}_2$ coincides with $\mathfrak{S}_1$ the splitting process is at an end. If not, we define $\mathfrak{S}_2 = \mathfrak{S}_1 - \mathfrak{R}_2$ and from $\mathfrak{S}_2$ we select a lowest element A_3. We have $\mu(A_3) > \mu(A_2)$. $\mathfrak{R}_3$ is defined in the obvious way. Continuing, we define a sequence $\mathfrak{S}_0, \mathfrak{S}_1, \mathfrak{S}_2, \cdots$, a sequence $A_1, A_2, A_3, \cdots$, and a sequence $\mathfrak{R}_0, \mathfrak{R}_1, \mathfrak{R}_2, \cdots$. For each i we have $\mathfrak{S}_i = \mathfrak{S}_{i-1} - \mathfrak{R}_i$. Each A_i is in $\mathfrak{R}_i$. The A_i constitute an ascending sequence; that is,

$\mu(A_1) < \mu(A_2) < \cdots$. There are no distinct subscripts j, k for which an element A exists such that $\mu(A_j) = \mu(AA_k)$.

12. We show that this process must terminate: that there must exist an n for which $\mathfrak{S}_{n-1} = \mathfrak{R}_n$. Suppose this were false, and, consequently, that the A_i constituted an infinite sequence. Assumption III would then imply that there was an integer m and elements $C_1, \cdots, C_m$ such that

$$\mu(A_{m+1} + C_1A_1 + \cdots + C_mA_m) < \mu(A_{m+1})$$

with

$$\mu(C_iA_i) \leq \mu(A_{m+1}), \qquad i = 1, \cdots, m.$$

Using I(e), we have $\mu(A_{m+1}) = \mu(C_1A_1 + \cdots + C_mA_m)$. By (4), then, we have $\mu(A_{m+1}) \leq \max(\mu(C_iA_i))$. These relations imply that for some j, $\mu(C_jA_j) = \mu(A_{m+1})$. This furnishes a contradiction, for the marks of the A_i cannot be so related. The finiteness of the various sequences is established.

13. We have shown that $\mathfrak{R}$ contains finitely many elements $A_1, \cdots, A_n$, which are such that the mark of any element of $\mathfrak{R}$ not a coefficient has a representation $\mu(CA_i)$. The marks of the A_i are connected by the relation $\mu(A_1) < \cdots < \mu(A_n)$. Observe that the mark of an element may have many representations of this sort, since several of the letters $A_1, \cdots, A_n$ may be eligible to serve as the A_i, and many elements may be admissible as the C for any given A_i.

It will now be shown that if C is any element of $\mathfrak{R}$ not a coefficient, then there is a power product of the A_i whose mark is the mark of C. There is a positive integer i_1, less than or equal to n, and an element C_1 of $\mathfrak{R}$ such that $\mu(C) = \mu(C_1A_{i_1})$. If C_1 is a coefficient, then $\mu(C_1A_{i_1}) = \mu(A_{i_1})$, and we have the desired representation for the mark of C. In any case $\mu(C_1) < \mu(C)$, since no A_i is a coefficient. If C_1 is not a coefficient its mark has a representation $\mu(C_2A_{i_2})$. Again we have $\mu(C_2) < \mu(C_1)$. If C_2 is a coefficient we have $\mu(C_1) = \mu(A_{i_2})$ and therefore, by I(d), $\mu(C) = \mu(A_{i_1}A_{i_2})$. Then $A_{i_1}A_{i_2}$ is the power product in question. If C_2 is not a coefficient this process can be continued. In this way a sequence of integers $i_1, i_2, \cdots$ and a sequence of elements $C_1, C_2, \cdots$ is determined, where $\mu(C_{k+1}) = \mu(C_kA_{i_k})$. Since $\mu(C_1) > \mu(C_2) > \cdots$, the sequences must be finite. Since the procedure can be applied for any C_i not a coefficient, it must be that one of the C_i, say C_t, is a coefficient. It follows that $\mu(C) = \mu(A_{i_1}A_{i_2} \cdots A_{i_t})$. This shows that the mark of C is the mark of a power product of the A_i.

14. We terminate this investigation of $\mathfrak{M}$ by showing that distinct power products of the A_i have distinct marks. For this purpose we need again the

fact that there is no valid relation $\mu(A_i) = \mu(AA_j)$ with i not equal to j. First we remark that distinct powers of any given A_i have different marks. This follows from the fact that an alleged relation $\mu(A_i^p) = \mu(A_i^{p+t})$ implies that A_i^t is a coefficient, whereas, by I(b), no power of an element which is not a coefficient can be a coefficient. This disposes of the case in which the set of elements A_i consists of the single element A_1.

15. Next we show that if an element A_2 exists, then, for all k, $\mu(A_1^k) < \mu(A_2)$. If this were not the case, then $\mu(A_1) < \mu(A_2) < \mu(A_1^{k+1})$. By assumption II there are elements P, Q, R for which

$$\mu(A_1 + PR) < \mu(A_1); \qquad \mu(A_2 + QR) < \mu(A_2); \qquad \mu(P) < \mu(A_1).$$

Recalling that A_1 is the lowest of those elements which are not coefficients, we see from the third relation that P is a coefficient. From the first we have, by I(e), that $\mu(A_1) = \mu(PR)$. This shows that P cannot be zero, and that $\mu(R) = \mu(A_1)$. From the second relation we have $\mu(A_2) = \mu(QR)$. It follows that $\mu(A_2) = \mu(QA_1)$. Since such a relation cannot be valid we have a contradiction. We proceed by induction, showing that if i is any positive integer less than n then the mark of any power product of $A_1, \cdots, A_i$ is less than that of A_{i+1}. Assuming the contrary, let there be such a power product of total degree k whose mark is not less than that of A_{i+1}. We have, *a fortiori*, $\mu(A_i) < \mu(A_{i+1}) < \mu(A_i^{k+1})$. Assumption II implies the existence of elements P, Q and R with the familiar properties. We have $\mu(A_i) = \mu(PR)$ and $\mu(P) < \mu(A_i)$. The inequality $\mu(P) < \mu(A_i)$ shows that the induction hypothesis applies to P. If P were not a coefficient its mark would be that of a power product of $A_1, \cdots, A_{i-1}$. The equation $\mu(A_i) = \mu(PR)$ would yield the impossible conclusion that the mark of A_i is the same as that of a multiple of A_j with j less than i. It must be, then, that P is a coefficient, whence $\mu(A_i) = \mu(R)$. This also leads to an impossibility, for the equation $\mu(A_{i+1}) = \mu(QR)$ would force the mark of A_{i+1} to be the mark of a multiple of A_i. This proves our statement.

16. The uniqueness proof is quickly completed. Let P and Q be two power products of $A_1, \cdots, A_n$ whose marks are identical. We are to show that the power products are identical. Assume that they are not. Neither of P, Q is a proper multiple of the other, for then we could not have $\mu(P) = \mu(Q)$. We may use (6) to cancel any factors common to P and Q. We obtain two power products P' and Q' which are relatively prime, of positive degree, and such that $\mu(P') = \mu(Q')$. One of these power products, say P', contains effectively an A_k whose subscript is larger than that of any A_i effectively

present in Q'. We have shown above that then $\mu(Q') < \mu(A_k)$. Certainly $\mu(Q') < \mu(P')$, which is a contradiction. This completes the uniqueness proof.

17. It is apparent from the above discussion how the power products of the A_i are ordered. If $P = A_1^{a_1} \cdots A_n^{a_n}$ and $Q = A_1^{b_1} \cdots A_n^{b_n}$ consider the smallest i for which $a_{n-i} - b_{n-i}$ differs from zero. If this difference is positive then $\mu(P) > \mu(Q)$. Of course if all the differences are zero then P and Q are identical.

The Generators of $\mathfrak{R}$.

18. We proceed to show that $\mathfrak{R}$ contains elements $X_1, \cdots, X_n$ which are such that $\mathfrak{R}$ is the ring $\mathfrak{R}_0[X_1, \cdots, X_n]$. We shall see that $\mu(X_i) = \mu(A_i)$, $i = 1, \cdots, n$.

19. Consider all those elements of $\mathfrak{R}$ whose marks do not exceed $\mu(A_1)$. Let this aggregate of elements be denoted by $\mathfrak{J}$. Let $\mathfrak{J}_1$ denote the subset of $\mathfrak{J}$ consisting of all elements of the form $LA_1 + S$; $L, S \in \mathfrak{R}_0$. If $\mathfrak{J}_1$ fills out $\mathfrak{J}$ then A_1 can serve as the element X_1. If $\mathfrak{J}$ contains elements not in $\mathfrak{J}_1$ let B_1 be such an element. We must have $\mu(B_1) = \mu(A_1)$ for otherwise B_1 would be a coefficient, and hence be included in $\mathfrak{J}$. Let $\mathfrak{J}_2$ denote the subset of $\mathfrak{J}$ consisting of all elements of the form $LA_1 + MB_1 + S$; $L, M, S \in \mathfrak{R}_0$. $\mathfrak{J}_2$ contains $\mathfrak{J}_1$ as a proper subset. If $\mathfrak{J}_2$ coincides with $\mathfrak{J}$ this process is at an end. If not we continue, selecting an element C_1 from $\mathfrak{J}$ which does not belong to $\mathfrak{J}_2$. Again $\mu(C_1) = \mu(A_1)$. Let $\mathfrak{J}_3$ be the totality of elements of the form $LA_1 + MB_1 + NC_1 + S$, $L, M, N, S \in \mathfrak{R}_0$. It is by now obvious how this procedure continues. We show that it must terminate after a finite number of steps, yielding a set of elements $A_1, B_1, C_1, \cdots, E_1$ such that $\mathfrak{J}$ coincides with the collection

$$LA_1 + MB_1 + NC_1 + \cdots + PE_1 + S; \quad L, M, N, \cdots, P, S \in \mathfrak{R}_0.$$

Suppose this were not the case. Then the set $A_1, B_1, C_1, \cdots$ would constitute an infinite sequence. No element of this sequence has a representation as the sum of an element of $\mathfrak{R}_0$ plus a linear combination of its predecessors with coefficients in $\mathfrak{R}_0$. Assumption III asserts that some element H_1 of this sequence must be such that its sum with a linear combination of its predecessors must have a mark less than that of H_1, that is, less than $\mu(A_1)$. Such a sum must be a coefficient. The assumption asserts further that each term of the linear combination is a product of the form DD_1 whose mark does not exceed $\mu(A_1)$. Since $\mu(D_1) = \mu(A_1)$, D is a coefficient. Thus H_1 actually has a representation of the kind it was supposed not to have. This establishes the finiteness of the sequence.

20. Let us apply assumption III to A_1 and B_1, denoting the three elements indicated in that assumption by P_1, Q_1, R_1. From $\mu(A_1 + P_1R_1) < \mu(A_1)$ it follows that $A_1 + P_1R_1$ is a coefficient. Similarly it follows from $\mu(B_1 + Q_1R_1) < \mu(B_1)$ and $\mu(P_1) < \mu(A_1)$ that $B_1 + Q_1R_1$ and P_1 are coefficients. On the other hand R_1 cannot be a coefficient; in fact from $\mu(A_1) = \mu(P_1R_1)$ we see that $\mu(A_1) = \mu(R_1)$. We have found an element R_1, which is such that both A_1 and B_1 have representations of the form $LR_1 + S$, $L, S \varepsilon \mathfrak{R}_0$. Treating R_1 and C_1 in the same way, we obtain an element R_2, such that both R_1 and C_1 have representations of the form $LR_2 + M$; $L, M \varepsilon \mathfrak{R}_0$. Of course A_1 and B_1 also have representations of this sort. Continuing in this way, we obtain an element X_1, which is such that each of the elements $A_1, B_1, \cdots, E_1$ has a representation of the form $LX_1 + M$; $L, M \varepsilon \mathfrak{R}_0$.

21. Consider, now, the totality of elements whose marks are less than $\mu(A_2)$. If there is no element A_2 the aggregate in question is $\mathfrak{R}$ itself. We shall show that this set is the polynomial ring $\mathfrak{R}_0[X_1]$.

Every element of this set which is not a coefficient has the same mark as some power of A_1, and, since $\mu(A_1) = \mu(X_1)$, has the same mark as some power of X_1. Let C be such an element. Then there is an integer c such that $\mu(C) = \mu(X_1^c)$. If c is unity we know that C belongs to $\mathfrak{R}_0[X_1]$, and, in fact, is of degree unity in X_1. We use an induction on c to prove that C likewise belongs to $\mathfrak{R}_0[X_1]$ when c exceeds unity, and that it is of degree c. In virtue of the relations $\mu(X_1) < \mu(C) < \mu(X_1^{c+1})$ we may use assumption II. We obtain thereby elements P, Q, R with the familiar properties. It follows from $\mu(P) < \mu(X_1)$ that P is a coefficient. From $\mu(X_1 + PR) < \mu(X_1)$ we see that $\mu(PR) = \mu(X_1)$, whence $\mu(R) = \mu(X_1)$. This last equation implies the existence of coefficients L and M for which $R = LX_1 + M$. Turning to the relation $\mu(C + QR) < \mu(C)$, we see that $\mu(C) = \mu(QR)$. Since R is not a coefficient, we must have $\mu(Q) < \mu(C)$. This shows that Q is subject to the induction hypothesis. We infer that it is an element of $\mathfrak{R}_0[X_1]$ whose degree is less than c. Similarly $C + QR$, because its mark is less than $\mu(C)$, is an element of $\mathfrak{R}_0[X_1]$ whose degree is less than c. Collecting our results relative to Q, R and $C + QR$, we see that C is an element of $\mathfrak{R}_0[X_1]$ whose degree does not exceed c. There remains to show only that the degree of C is c. To do this we apply (4) to C, using for C its representation $C_0 + C_1X_1 + \cdots + C_cX_1^c$, where the C_i are in $\mathfrak{R}_0$. It cannot be that $C_c = 0$ for then the mark of C could not be $\mu(X_1^c)$. This shows that the degree of C is c.

22. If the set $A_1, \cdots, A_n$ consists of the single element A_1 then the

assertion that $\mathfrak{R}$ is a polynomial ring is verified. If this is not the case we still have the problem of selecting elements $X_2, \cdots, X_n$ such that R is the polynomial ring $\mathfrak{R}_0[X_1, X_2, \cdots, X_n]$. Assuming that n exceeds unity we show that $\mathfrak{R}$ contains an element X_2, which is such that the totality of elements of $\mathfrak{R}$ whose marks are less than $\mu(A_3)$ is the ring $\mathfrak{R}_0[X_1, X_2]$. (We understand this to mean that if $n = 2$ then $\mathfrak{R}$ itself is $\mathfrak{R}_0[X_1, X_2]$). We shall see that $\mu(X_2) = \mu(A_2)$. In selecting X_2 we follow **19**. Corresponding to the $\mathfrak{I}$ used there, we define $\mathfrak{U}$ to be the totality of elements of $\mathfrak{R}$ whose marks do not exceed $\mu(A_2)$. Let $\mathfrak{U}_1$ be the set of elements which have representations $LA_2 + S$, with S in $\mathfrak{R}_0[X_1]$ and L in $\mathfrak{R}_0$. $\mathfrak{U}_1$ is contained in $\mathfrak{U}$. If it coincides with $\mathfrak{U}$ then A_2 can serve as the element X_2. In the contrary case $\mathfrak{U}$ contains an element B_2 not in $\mathfrak{U}_1$. The mark of B_2 must be $\mu(A_2)$, for otherwise it would be less than this mark, forcing B_2 to be in $\mathfrak{U}_1$. Let $\mathfrak{U}_2$ be the totality of elements which have representations $LA_2 + MB_2 + S$; $L, M\ \mathfrak{R}_0$, $S \in \mathfrak{R}_0[X_1]$. If $\mathfrak{U}_2$ does not coincide with $\mathfrak{U}$ we continue as in **19** obtaining a sequence $A_2, B_2, C_2, \cdots$. No element of this sequence is a sum of an element of $\mathfrak{R}_0[X_1]$ plus a linear combination of its predecessors with coefficients in $\mathfrak{R}_0$. The mark of each element of this sequence is $\mu(A_2)$. These facts make it possible to repeat the reasoning of **19**, showing that this sequence must be finite. There is only one variation which occurs in the proof. There frequent use was made of the inference that an element which is lower than A_1 must be in $\mathfrak{R}_0$. Here the corresponding step is that an element whose mark is less than $\mu(A_2)$ must be in $\mathfrak{R}_0[X_1]$.

23. By using the reasoning of **20** with the modifications described above, we find an element X_2 with the following properties. The mark of X_2 is $\mu(A_2)$. Every element of $\mathfrak{R}$ which has this mark has a representation $LX_2 + M$, $L \in \mathfrak{R}_0$, $M \in \mathfrak{R}_0[X_1]$.

24. Let C be any element not in $\mathfrak{R}_0[X_1]$ which is lower than A_3. Its mark is that of a power product of X_1, X_2 which involves X_2 effectively. Let this power product be $X_1^{c_1}X_2^{c_2}$, c_2 being different from zero. We show that C is in $\mathfrak{R}_0[X_1, X_2]$ and that it has a representation as a sum $KX_1^{c_1}X_2^{c_2} + \Sigma K_iP_i$, where K and the K_i are in $\mathfrak{R}_0$, K is not zero, and the P_i are power products of X_1, X_2 all lower than $X_1^{c_1}X_2^{c_2}$.

We use a contradiction proof, assuming that such an element C exists which does not have the properties indicated. We assume that C is the lowest of all elements which satisfy the hypothesis of our statement, but not its conclusion. It is permissible to assume that a lowest element (possibly several) exists because $\mathfrak{M}$ is well-ordered. Note that for such a C, $\mu(C) > \mu(X_2)$ because if $\mu(C) \leqq \mu(X_2)$, C has a representation $LX_2 + M$, $L \in \mathfrak{M}_0$, $M \in \mathfrak{R}_0[X_1]$.

We have $\mu(X_2) < \mu(C) < \mu(X_2^{c_2+1})$. We apply assumption II, obtaining elements P, Q, R. It follows from $\mu(P) < \mu(X_2)$ that P is in $\mathfrak{R}_0[X_1]$, and consequently that the mark of P is either $\mu(1)$ or that of a power of X_1. It follows from $\mu(X_2 + PR) < \mu(X_2)$ that $\mu(X_2) = \mu(PR)$. Thus P must be a coefficient, for if its mark were that of a power of X_1 this last equation would express the mark of X_2 as that of a multiple of X_1. Since P is a coefficient, we have $\mu(R) = \mu(X_2)$, whence R has a representation $LX_2 + M$, $L \in \mathfrak{R}_0$, $M \in \mathfrak{R}_0[X_1]$. Turning to the relation $\mu(C + QR) < \mu(C)$ we infer that $\mu(C) = \mu(QR)$. It follows from (6) that the mark of Q is $\mu(X_1^{c_1}X_2^{c_2-1})$. Because of the minimal character of C, Q has a representation $K'X_1^{c_1}X_2^{c_2-1} + \Sigma K'_iP'_i$, where K' and the K'_i are in $\mathfrak{R}_0$, K' is different from zero and the P'_i are power products of X_1, X_2 all lower than $X_1^{c_1}X_2^{c_2-1}$. Evaluating the product QR, we see that it has a representation $K'LX_1^{c_1}X_2^{c_2} + \Sigma K''_j P''_j$ the character of the elements K''_j, P_j being obvious. We attain our contradiction by observing that since $C + PR$ is lower than C, it has a representation as an element of $\mathfrak{R}_0[X_1, X_2]$, no monomial of which is as high as $X_1^{c_1}X_2^{c_2}$. We have $C = (C + PR) - PR$; the analysis of $C + PR$ and of PR makes it clear that C has a representation of the kind described. Thus the assumption that there exists an element lower than A_3 which does not have such a representation must be discarded. This shows that the totality of elements lower than A_3 is the ring $\mathfrak{R}_0[X_1, X_2]$.

25. It is by now plain how the proof can be conducted so that finally a ring $\mathfrak{R}_0[X_1, \cdots, X_n]$ is obtained which coincides with $\mathfrak{R}$. The connection between the mark of an element of $\mathfrak{R}$ and its representation as a polynomial in $X_1, \cdots, X_n$ is also obvious. This verifies the statements made concerning the first list of assumptions.

The Alternate List of Assumptions.

26. We now replace assumptions II and III by II′, III′ and IV′. We show that the second list implies the first, so that all of the foregoing can be retained. We show further that, in addition to the above limitations imposed upon $\mathfrak{R}$, the new assumptions imply that $\mathfrak{R}_0$ is a field.

27. Let us show that II′, III′, IV′ imply III. Consider any infinite sequence $A_1, A_2, \cdots$, of non-zero elements. By III′ there are integers i and j and an element C such that $\mu(A_{i+j}) = \mu(CA_i)$. By IV′ there is an element C' such that $\mu(A_{i+j} + C'CA_i) < \mu(A_{i+j})$. This covers the first part of III, the integer m being $i + j$ and the elements C_k being all zero except C_i which is $C'C$. We have only to show that $\mu(C'CA_i) \leq \mu(A_{i+j})$. This is a consequence of I(e); actually we know that $\mu(A_{i+j} + C'CA_i) < \mu(A_{i+j})$ implies $\mu(A_{i+j}) = \mu(C'CA_i)$.

28. To show that II′, III′, IV′ imply II we shall need several results of the early part of this paper. They are that $\mathfrak{M}$ is well-ordered, that $\mu(A) = \mu(-A)$ and that $\mathfrak{R}$ contains a unit. These assertions were all verified in **4-6**, without the use of II, so that the results can properly be used in the present context without further argument.

Assumption II has two cases. The first concerns elements A, B and an integer m for which $\mu(A) = \mu(B) < \mu(A^m)$. Since $\mu(A) < \mu(A^m)$ we see that A is not zero. Applying IV′ to A and B we find a C such that $\mu(A + CB) < \mu(A)$. It is now possible to deduce the conclusions of assumption II, the elements P, Q, R being respectively C, -1, B. We already have $\mu(A + PR) < \mu(A)$. The relation $B + (-1)B = 0$ would verify $\mu(B + QR) < \mu(B)$ if it were known that $\mu(0) < \mu(B)$. This follows immediately from the relations $\mu(A) = \mu(B)$ and $A \neq 0$. There remains to show that $\mu(C) < \mu(A)$. We have $\mu(A) = \mu(CB)$, so that $\mu(C) \leq \mu(A)$. If $\mu(C)$ were equal to $\mu(A)$ we should then have $\mu(A) = \mu(AB)$, so that $\mu(B)$ would be equal to $\mu(1)$. This cannot be the case, for $\mu(A^m) = \mu(B^m) > \mu(B)$, whereas $\mu(1^m) = \mu(1)$. This covers the first case of assumption II.

In the second case we have $\mu(A) < \mu(B) < \mu(A^m)$. Assumption II′ applies, yielding elements P', Q', R' such that $\mu(A) = \mu(P'R')$, $\mu(B) = \mu(Q'R')$, $\mu(P') < \mu(A)$. Applying IV′ to the two equations just obtained, we find elements C and D for which $\mu(A + CP'R') < \mu(A)$, $\mu(B + DQ'R') < \mu(B)$. To deduce assumption II from these results let the elements P, Q, R be, respectively, CP', DQ', R'. All that need be shown to complete the discussion of this case is that $\mu(CP') < \mu(A)$. It follows from $\mu(A + CP'R') < \mu(A)$ that $\mu(A) = \mu(CP'R')$, and, consequently, that $\mu(CP') \leq \mu(A)$. If $\mu(CP')$ were equal to $\mu(A)$ it would follow that $\mu(R')$ would be equal to $\mu(1)$. But if $\mu(R')$ were equal to $\mu(1)$ we should have $\mu(R'P') = \mu(P')$ and hence $\mu(P') = \mu(A)$. This contradiction shows that $\mu(CP')$ is in fact less than $\mu(A)$, completing the proof.

29. It is now proper to discuss $\mathfrak{R}_0$ in connection with the second set of assumptions, because the assumptions of this set have been shown to imply those of the first set. Let us show that the ring $\mathfrak{R}_0$ obtained with the second list must be a field.

Let A and B be any non-zero elements of $\mathfrak{R}_0$. We have $\mu(A) = \mu(B)$ and can apply IV′. We obtain an element C such that $\mu(A + CB) < \mu(A)$. It follows that $A + CB = 0$, since no non-zero element of $\mathfrak{R}$ is lower than a coefficient. Thus the equation $A + XB = 0$ has a solution in $\mathfrak{R}$. We need only show that its solution C is in $\mathfrak{R}_0$. This follows readily. From $\mu(A + CB) < \mu(A)$ we see that $\mu(A) = \mu(CB)$. C cannot be higher than A, for if it were CB would be either higher than A or zero, and it is neither. From $\mu(C) = \mu(A)$ we therefore can conclude that C is in $\mathfrak{R}_0$.

Independence of the Assumptions.

30. We shall not embark upon a formal discussion of the independence of the assumptions. We present instead a few indications of the role of some of the assumptions by suppressing them and exhibiting counter-examples. In every instance we assume that I(a)-I(e) hold, and suppress only the later assumptions. In constructing the examples we use the ring $\mathfrak{k}$ of rational integers, and the field $\mathfrak{K}$ of rational numbers.

(i) III *is independent of* I(a)-I(e), II. Let $X_1, X_2, \cdots$ be an infinite sequence of indeterminates and consider the ring $\mathfrak{k}[X_1, X_2, \cdots]$. The elements of this ring may be ordered according to the plan set forth in 2. When its elements are so ordered the ring obeys assumptions I(a)-I(e), II. It does not satisfy III; the sequence $X_1, X_2, \cdots$ contains no elements corresponding to the A_{m+1} of that assumption.

(ii) II *is independent of* I(a)-I(e), III. Let $\mathfrak{R}$ be the ring $\mathfrak{k}[X, Y]$, where X and Y are indeterminates. Let the ordered set $\mathfrak{M}$ consist of $-\infty$, $0, 1, 2, 3, \cdots$, ordered in the obvious way. Let the mark of an element of $\mathfrak{R}$ be its total degree in X and Y, using $-\infty$ for the mark of zero, and 0 for the mark of the other elements of $\mathfrak{k}$. Then $\mathfrak{R}$ satisfies I(a)-I(e), III. It does not satisfy II, for while $\mu(X) = \mu(Y) < \mu(X^2)$ there are no elements P, Q, R of the sort described in that assumption. Observe that $\mathfrak{R}$ is a polynomial ring of precisely the kind we have discussed. The exceptional character of $\mathfrak{R}$ is to be attributed to the way it is ordered rather than to one of its internal features.

(iii) II′ *is independent of* I(a)-I(e), III′, IV′. Let $\mathfrak{R}$ be the ring $\mathfrak{K}[X, Y]$, where X and Y are indeterminates. Let $\mathfrak{M}$ be the set $-\infty$, $a + b2^{\frac{1}{2}}$, where a and b can assume all non-negative integral values. The elements of $\mathfrak{M}$ are ordered according to numerical value. Let the mark of zero be $-\infty$, that of a non-zero element of $\mathfrak{K}$ be 0, and that of a term X^aY^b be $a + b2^{\frac{1}{2}}$. If the mark of an arbitrary element of $\mathfrak{R}$ not in $\mathfrak{K}$ is defined to be the greatest of the marks of terms effectively present in it, then $\mathfrak{R}$ satisfies I(a)-I(e), III′, IV′. It does not satisfy II′, for $\mu(X) < \mu(Y) < \mu(X^2)$, and $\mathfrak{R}$ contains no elements P, Q, R as required by II′.

(iv) III′ *is independent of* I(a)-I(e), II′, IV′. The ring of example (i), with $\mathfrak{K}$ replacing $\mathfrak{k}$, furnishes an illustration of this.

(v) IV′ *is independent of* I(a)-I(e), II′, III′. Any polynomial ring which satisfies all the assumptions of our first set and whose coefficients do not constitute a field is an example which proves the independence of IV′.

Columbia University.

PROJECTIVE GEOMETRIES AS MULTIGROUPS.*

By WALTER PRENOWITZ.

1. Introduction. Projective geometry is traditionally called the geometry of projection and section. It studies the operations *projection* (*join*) and *section* (*meet*) in the set of *linear spaces* (points, lines, planes etc.). Different treatments of projective geometry introduce these ideas in different ways.

Veblen and Young [1] in their classic text take the terms *point* and *line* as primitive and characterize them by the following postulates:

PG1. *A line is a set of points.*

PG2. *Two distinct points belong to one and only one line.*

PG3. *If points a, b, c do not colline and p, q are distinct points such that a, b, p colline, and b, c, q colline then there exists a point r such that c, a, r colline and p, q, r colline.*

PG4. *Every line contains at least three points.*[1]

They introduce linear spaces of higher (finite) dimension inductively and define projection and section piecemeal for different elements. Their proofs of the basic alignment theorems for linear spaces lean heavily on PG3, *the triangle postulate,* and are strongly motivated by geometrical intuition.

The traditional description of projective geometry suggests a different axiomatization of the subject with *linear space, join* and *meet* as the primitive ideas. This has been done (in somewhat different ways) by Garrett Birkhoff [2] and K. Menger,[3] working independently. In their work, the point is not singled out as a building block for the construction of linear spaces of higher dimension. Linear spaces are given as undifferentiated elements or "chunks" which con-

* Received January 7, 1942. Presented in part to the American Mathematical Society, October 1939, under the title "Projective geometry as a group-like system."

[1] We have restated assumptions A, E_0 of Veblen and Young [1] pp. 16, 18 and added PG1 which they assume tacitly. We omit further extension and closure postulates so that a projective geometry need *not* be of finite dimension.

[2] See Birkhoff [1], [2]. Birkhoff's work has resulted in the elegant characterization of finite-dimensional projective geometries as *complemented modular lattices of finite dimensions which are simple* (Birkhoff [2], pp. 67-68).

[3] See Menger [1], [2]. The latter contains further references to work of Menger and his students on this topic. Menger has also axiomatized affine and other geometries in a similar way. See [2]; "A new foundation of non-Euclidean, affine, real projective and Euclidean geometry," *Proceedings of the National Academy of Science,* vol. 24 (1938), pp. 486-90; "Non-Euclidean geometry of joining and intersecting," *Bulletin of the American Mathematical Society,* vol. 44 (1938), pp. 821-4.

stitute a domain of operation for join and meet. The postulates are general properties of join and meet which permit a deduction of theorems by algebraic methods unaided by geometric intuition.

We shall try to develop projective geometry so as to attain both the intuitive simplicity of the Veblen and Young approach and the elegance and generality of the algebraic methods of Birkhoff and Menger. On the one hand we choose *point* as a primitive idea. Then, instead of following the usual development, let us try to introduce join and meet as quickly as possible and study their general properties by algebraic methods. First we observe that any figure or space (e. g. line, plane, quadric surface) may be considered to be a set of points. Hence the meet of two figures need not be a primitive idea but may be *defined* as their *intersection* or *set theoretic product.* Secondly, *the join of any two figures may be obtained by a "criss-cross" process of joining each point of the first figure to each point of the second and combining into a set all points on all "joins" formed in this way.*[4] Thus it should be possible to base projective geometry on the idea *point* and the operation *join of a point and a point.*

These considerations suggest a study of the following abstract system. Consider a set G of elements $a, b, c, \cdots$ which may be called *points* and a many-valued two-term operation $+$ which associates to each pair a, b a set $a+b$ called the *join* or *sum* of a and b. Concretely interpreted, $a+b$ is the set of points on the "line" ab if $a \neq b$, and $a+a$ is just a. We postulate J1, $\cdots$, J6.

J1. *If* $a, b \subset$[5] G, $a+b$ *is a uniquely determined subset of* G.

J2. *If* $a, b \subset G$, $a+b \supset a, b$.

J3. *If* $a, b \subset G$, $a+b = b+a$.

J4. *If* $a \subset G$, $a+a = a$.[6]

In order to state the associative law we need

DEFINITION 1. *If sets* $A, B \subset G$ *and* $A, B \neq O$, *the null set, the join or sum of* A *and* B, *denoted* $A+B$, *is the set theoretic sum* $\sum_{a \subset A,\, b \subset B} (a+b)$.[7] This is

[4] In this sense for example a conical surface is the join of a point and a curve and a projective 3-space is the join of two skew lines.

[5] We use the inclusion signs $\supset$, $\subset$ both for sets and elements and do not distinguish between the element a and the set (a) whose only element is a.

[6] There appears to be a difficulty here since J1 postulates that $a+a$ is a set and J4 that it is an element. There is no inconsistency since we have agreed to identify a with (a)—see the preceding footnote.

[7] Observe that A or B may be single elements and that if both are, say $A = a$, $B = b$ then $A+B$ as defined is $a+b$. Thus no inconsistency arises between the defined term $+$ and the primitive term $+$.

merely the formal description in our system of the "criss-cross" process described above, and is the natural analogue of the notion, product of two sets, in ordinary group theory. To complete our definition we take $A + O = O + A = A$ for any $A \subset G$.

J5. *If* $a, b, c \subset G$, $(a + b) + c = a + (b + c)$.

J6. *If* $a_1, a_2, b \subset G$ *and* $a_1 + a_2 \supset b \neq a_1$ *then* $a_1 + a_2 = a_1 + b$.

J1, · · ·, J6 are easily seen to be verified in any projective geometry (i.e. any system satisfying PG1, · · ·, PG4). For J1, · · ·, J4 this is trivial. J5 is essentially equivalent to PG3 but has greater deductive power since it does not involve the restriction that a, b, c be non-collinear. J6 is a weak form of PG2. Examining the postulates from an algebraic viewpoint we observe that J1 is a closure postulate for the many-valued operation $+$, and J3, J5 are perfectly familiar. J4 appears in Boolean algebra and lattice theory. J6, in a sense the least "algebraic" of the postulates, may be characterized as an "exchange" [8] property since it enables us to exchange a_2 for b without altering the sum $a_1 + a_2$.[9]

The system $(G, +)$ characterized by J1, · · ·, J6 is group-like in character. It is in fact a special case of the generalized group in which multiplication is many-valued, studied in recent years by Marty, Wall, Ore, Dresher and others under the names *hypergroup* or *multigroup*.[10] We propose to show that the basic ideas of projective geometry of *finite or infinite dimension* are naturally expressed and studied by *group theoretic* concepts. We shall develop our results directly from J1, · · ·, J6 without appeal to the general theory of multigroups or to lattice theory, making the paper complete in itself.

2. Relation to the system of Veblen and Young. As we have noted, J1, · · ·, J6 are implied by PG1, · · ·, PG4. In order to consider the converse question we must define *line* in our system. Thus we adopt

DEFINITION 2. *If* $a, \neq b \subset G$ *we call* $a + b$ *a line.*

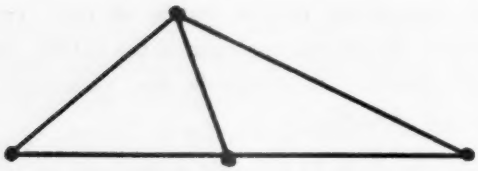
It is then easy to show that J1, · · ·, J6 *imply* PG1, PG2, PG3. They do *not* imply PG4 since there exist systems satisfying J1, · · ·, J6 but not PG4. An example is indicated in the adjoining diagram.

---

[8] Cf. van der Waerden [1], p. 96 Satz 4.

[9] J6 is equivalent in the presence of J1, · · ·, J5 to the following "transposition" principle: *If* $a_1 + a_2 \supset b \neq a_1$, *then* $a_1 + b \supset a_2$. Cf. MacLane [1] axioms (E_1), (E_2); Menger [2], p. 461 the assertion equivalent to Law + 6.

[10] This was pointed out to the author by Dr. H. H. Campaigne. A list of references on multigroups is given in Dresher and Ore, "Theory of multigroups," *American Journal of Mathematics*, vol. 60 (1938), pp. 705-733.



However if we replace PG4 by the weaker postulate

PG4′. *Every line contains at least two points,*

we can easily show that PG1, · · ·, PG4′ and J1, · · ·, J6 are equivalent.[11]

Although our postulate system is weaker than that of Veblen and Young, we shall not strengthen it now since the basic alignment for linear spaces is deducible from it. Later (Sections 9, 10) we shall characterize classic projective geometries as systems satisfying our postulates.

3. Preliminary properties of join. We proceed to study systems $(G, +)$ satisfying J1, · · ·, J6. To emphasize in general the analogy with classical abelian group theory and in particular the closure of G under operation $+$, we shall say that G is a *group with respect to* $+$, or simply that G is a *group*.[12] We shall consider closure under the inverse operation later. We denote elements of G by small letters $a, b, c, \cdots$ and subsets of G by capitals $A, B, C, \cdots$.

We derive some elementary results on sums of sets analogous to those in classical group theory for products of sets. We have directly from Definition 1

THEOREM 1. *If* $x \subset a + b$, $a \subset A$, $b \subset B$ *then* $x \subset A + B$. *Conversely if* $A, B \neq O$ *and* $x \subset A + B$ *then* $x \subset a + b$ *where* $a \subset A$, $b \subset B$.

We generalize J2 for sets A, B in

THEOREM 2. $A + B \supset A, B$.

Proof. If A or $B = O$ this is trivial. If $A, B \neq O$ let $a \subset A$, $b \subset B$. Then by J2, $a \subset a + b$ so that $a \subset A + B$ by Theorem 1. Thus $A \subset A + B$. Similarly $B \subset A + B$.

Now we derive a simple but very useful "monotonic" law.

THEOREM 3. *If* $A \subset B$, $C \subset D$ *then* $A + C \subset B + D$.

Proof. If $A = O$ or $C = O$ the result follows from Theoren 2. Suppose

---

[11] We might call systems satisfying PG1, · · ·, PG4′ *generalized projective geometries.*

[12] To avoid ambiguity we shall refer to groups of the ordinary kind, where the composition is single-valued, as *classical groups.*

$A, C \neq O$. Let $x \subset A + C$. Then by Theorem 1, $x \subset a + c$ where $a \subset A$, $c \subset C$. In view of our hypothesis $a \subset B$, $c \subset D$ so that $x \subset B + D$ by Theorem 1. Thus $A + C \subset B + D$.[13]

COROLLARY. *If* $A \subset B$ *then* $A + X \subset B + X$.

We generalize J3, J5 in

THEOREM 4. (a) $A + B = B + A$; (b) $(A + B) + C = A + (B + C)$.

Proof. (a) This follows readily from J3.

(b) If A, B or C is O the result holds by definition. Suppose $A, B, C \neq O$. Let $x \subset (A + B) + C$. Then $A + B \neq O$ so that by Theorem 1, $x \subset p + c$ where $p \subset A + B$, $c \subset C$. Likewise by Theorem 1, $p \subset a + b$ where $a \subset A$, $b \subset B$. By Theorem 3, $b + c \subset B + C$. Hence by Theorem 3 and J5

$$x \subset p + c \subset (a + b) + c = a + (b + c) \subset A + (B + C)$$

so that $(A + B) + C \subset A + (B + C)$. Similarly we show $A + (B + C) \subset (A + B) + C$.

We define *finite sums* by recursion.

DEFINITION 3. $\sum_{i=1}^{1} A_i = A_1, \quad \sum_{i=1}^{n+1} A_i = \sum_{i=1}^{n} A_i + A_{n+1}$.[14]

By induction from Theorems 2, 3, 4 we derive

THEOREM 5. (a) *The associative and commutative laws hold for finite sums* $A_1 + \cdots + A_n$;

(b) $A_1 + \cdots + A_n \supset A_i$, $1 \leqq i \leqq n$;

(c) $A_i \subset B_i$, $1 \leqq i \leqq n$, *implies* $A_1 + \cdots + A_n \subset B_1 + \cdots + B_n$.[15]

4. Linear spaces. The principal object of study in projective geometry is the *linear space*. We naturally characterize linear spaces by the intuitively familiar property of *linearity* or *flatness—if a linear space contains two distinct points, it contains the line joining them*. This suggests

---

[13] Observe that this property (and all others involving sums of arbitrary sets) holds for *individual elements*—we might have for example $A = a$ and $C = c$.

[14] If $m \leqq n$ we adopt the usual conventions for $\sum_{i=m}^{n} A_i$ and if $m > n$ we take this to be O.

[15] In general in this paper the subscript n denotes a *non-negative* integer.

DEFINITION 4. *$A \subset G$ is called a linear subspace of G or simply a linear space, if $A \supset x, y$ implies $A \supset x + y$.*[16]

Thus *A is a linear space if and only if it is closed under $+$.* This suggests the classical idea of *subgroup.* Let us say $A \subset G$ is a *subgroup* of G if A is a group with respect to $+$, i.e. if A satisfies J1, $\cdots$, J6 for the addition operation of G. Clearly *A is a subgroup of G if and only if A is closed under $+$.* Thus *A is a linear space if and only if it is a subgroup of G* and hereafter we shall use the terms *linear space* and *subgroup* (or *group*) interchangeably.

We proceed to derive some elementary properties of linear spaces.

THEOREM 6. *A is a linear space (subgroup of G) if and only if $A + A = A$.*

Proof. Let A be a linear space. If $A = O$ certainly $A + A = A$. If $A \neq O$, $A \supset A + A$ since A is closed under $+$. By Theorem 2, $A + A \supset A$. Thus $A + A = A$. Conversely $A + A = A$ implies $A \supset A + A$ so that A is closed under $+$ and is a linear space.

Now we can easily derive the following "absorption" law.

COROLLARY 1. (Absorption) *If a linear space $A \supset X$ then $A + X = A$.*

Proof. Since $A \supset X$ we have by Theorem 6 and the corollary to Theorem 3, $A = A + A \supset A + X$. By Theorem 2, $A + X \supset A$. Thus $A + X = A$.

COROLLARY 2. *If a linear space $A \supset X_1, \cdots, X_n$ then $A \supset X_1 + \cdots + X_n$.*

Proof. We have $A \supset X_i$, $1 \leqq i \leqq n$. Adding these relations as "inequalities" by Theorem 5(c) and applying Theorem 6 we get the desired result.

We now derive a property which is sometimes used to define the join of two linear spaces.

THEOREM 7. *If A, B are linear spaces (subgroups of G) $A + B$ is the least linear space (subgroup of G) containing A and B.*

Proof. By Theorems 5(a) and 6

$$(A + B) + (A + B) = (A + A) + (B + B) = A + B.$$

Thus $A + B$ is a linear space by Theorem 6. By Theorem 2, $A + B \supset A, B$. Finally, if a linear space $X \supset A, B$ then $X \supset A + B$ by the last corollary.

Using a similar argument, or by induction, we prove

---

[16] Observe that G, O and individual elements $a, b, c, \ldots$ are linear spaces. O plays a role similar to that of the *identity group* in classic group theory.

COROLLARY 1. *If $A_1, \cdots, A_n$ are linear spaces $A_1 + \cdots + A_n$ is the least linear space containing $A_1, \cdots, A_n$.*

Taking the A's to be individual elements we have

COROLLARY 2. *$a_1 + \cdots + a_n$ is the least linear space containing $a_1, \cdots, a_n$.*[17]

5. Determination of linear spaces. We now consider the idea, *linear space determined by a set of elements.* For example, in projective geometry we speak of the *line l determined by points a, b.* We can take this to mean that l is the "*simplest*" or *least* linear space which contains a and b. This suggests

DEFINITION 5. *Let $S \subset G$. By the linear space (subgroup) determined or generated by S, denoted $\{S\}$, we mean the least linear subspace of G which contains S.*[18] *We say S is, or its elements constitute, a set of generators of $\{S\}$. If S is finite* and its elements are s_i, $1 \leqq i \leqq n$, where $n \geqq 0$ and the s_i are not necessarily distinct, we find it convenient to write $\{S\} = \{s_1, \cdots, s_n\}$. Observe that $\{O\} = O$ so that $\{s_1, \cdots, s_n\} = O$ if $n = 0$.

The existence and uniqueness of $\{S\}$ is shown exactly as in classical group theory by considering the set of all linear spaces (subgroups of G) which contain S. Thus we may assert

THEOREM 8. *For each $S \subset G$, $\{S\}$ exists and is uniquely determined.*

The next theorem asserts that $\{S\}$, the group generated by S, can be constructed essentially as in classical *finite* group theory. The proof is left to the reader.

THEOREM 9. *$\{S\}$ consists of all elements x of G which are contained in finite sums of elements of S.*

COROLLARY 1. *If $S \supset T$ then $\{S\} \supset \{T\}$.*

COROLLARY 2. *$\{S \cup T\} = \{S\} + \{T\}$.*[19]

Proof. Let $x \subset \{S \cup T\}$. Then by Theorem 9, $x \subset S' + T'$ where S', T' are finite sums of elements of S, T respectively. Thus by Theorem 9, $S' \subset \{S\}$,

---

[17] That $a_1 + \cdots + a_n$ is a linear space is essentially Theorem $S_n 1(b)$ Veblen and Young [1], p. 30, that *an n-space contains the line joining any two of its points.* Compare their proof.

[18] This is the exact analogue of the idea, *group generated by a set of elements,* in classic group theory.

[19] $S \cup T$ denotes the *set union* or *logical sum* of sets S, T.

$T' \subset \{T\}$ so that by Theorem 3, $x \subset \{S\} + \{T\}$ and $\{S \cup T\} \subseteq \{S\} + \{T\}$. The converse inclusion is proved similarly.

When S is *finite* we get a neater form for Theorem 9 by restating Corollary 2 of Theorem 7:

THEOREM 10. $\{a_1, \cdots, a_n\} = a_1 + \cdots + a_n$.[20]

6. Linear independence. In forming a set of generators for a given linear space we naturally try to eliminate superfluous elements. This suggests

DEFINITION 6. *Let $S \subset G$. Suppose $\{S \dot{-} x\}$ [21] $\supset x$ for each $x \subset S$. Then we call S a linearly independent set or simply an independent set.*[22,23]

From this definition it is easy to derive the following theorems.

THEOREM 11. *Any subset of an independent set is also independent.*

THEOREM 12. *Let S be a set of generators of group G. Then S is independent if and only if S is a minimal set of generators of G.*

As in classical group theory we prefer a set of generators to be independent. This suggests

DEFINITION 7. *An independent set of generators of a group G is called a basis of G.*

Let a group G have a *finite* set of generators S. Deleting superfluous elements from S, one by one, we obtain finally a *minimal* set of generators of G which is a basis of G by Theorem 12. Thus we have proved the

COROLLARY. *Any group with a finite set of generators has a basis.*

We have not yet employed J6. We use it now to derive two lemmas which are generalizations of J6 and which lead rapidly to several important results in the theory of independence including Theorem 14, the *exchange theorem.*

LEMMA 1. *If $\{A \cup x\} \supset b$ but $\{A\} \not\supset b$ we may exchange x for b without altering $\{A \cup x\}$,* i. e. *$\{A \cup x\} = \{A \cup b\}$.*[24]

---

[20] Compare the corresponding result in classical abelian group theory that *the group generated by $a_1, \cdots, a_n$ is the set of all elements of the form $a_1 a_1 + \cdots + a_n a_n$ where the a's are integers.*

[21] We use the symbol $\dot{-}$ to denote *set theoretic* subtraction.

[22] For finite S this is essentially the definition adopted by Menger [2].

[23] Observe that O or a set consisting of a single element is independent.

[24] From this we can easily derive the exchange axiom (E_1) of MacLane [1]. Then if we wish to introduce lattice theoretic methods we can formulate our system as an

Proof. If $A = O$ then $x = b$ and the lemma is true. Suppose $A \neq O$. By hypothesis and Corollary 2 of Theorem 9, $b \subset \{A \cup x\} = \{A\} + x$. Hence we have

$$b \subset a + x \tag{1}$$

where $a \subset \{A\}$. $b \neq a$ since by hypothesis $b \not\subset \{A\}$. Thus we may apply J6 to (1) getting $a + x = a + b$. Adding $\{A\}$ to both members and "absorbing" a into $\{A\}$ (Corollary 1 of Theorem 6), we have $\{A\} + x = \{A\} + b$ so that by Corollary 2 of Theorem 9 $\{A \cup x\} = \{A \cup b\}$.

LEMMA 2. *If* $\{a_1, \cdots, a_m\} \supset b$ *but* $\{a_1, \cdots, a_p\} \not\supset b$ *for some* p, $0 \leqq p \leqq m$, *there is an* a_i, $p < i \leqq m$, *which may be exchanged for* b *without altering* $\{a_1, \cdots, a_m\}$, i. e. $\{a_1, \cdots, a_m\} = \{a_1, \cdots, a_{i-1}, b, a_{i+1}, \cdots, a_m\}$.

Proof. Let i be the *least* index for which $\{a_1, \cdots, a_i\} \supset b$. Then $p < i \leqq m$ and $\{a_1, \cdots, a_{i-1}\} \not\supset b$. Thus by Lemma 1, taking A to be the set composed of $a_1, \cdots, a_{i-1}$ we get

$$\{a_1, \cdots, a_i\} = \{a_1, \cdots, a_{i-1}, b\}$$

and adjoining $a_{i+1}, \cdots, a_m$ to the terms in each member we have the desired result.

COROLLARY. *If* $\{A\} \supset b$ *then* $\{A\} = \{A' \cup b\}$ *where* A' *is a proper subset of* A.[25]

We now derive a criterion for the independence of $S \cup T$.

THEOREM 13. *Let* $S \cdot T$ [26] $= O$. *Then* $S \cup T$ *is an independent set if and only if* S, T *are independent and* $\{S\} \cdot \{T\} = O$.

Proof. Suppose $S \cdot T = O$, $S \cup T$ independent. Then S, T are independent by Theorem 11. Suppose $\{S\} \cdot \{T\} \neq O$. Let $x \subset \{S\}, \{T\}$. By the last corollary $\{S\} = \{S' \cup x\}$ where S' is a proper subset of S. Hence by Corollary 2 of Theorem 9 and Corollary 1 of Theorem 6

$$\begin{aligned} \{S \cup T\} &= \{S\} + \{T\} = \{S' \cup x\} + \{T\} \\ &= \{S'\} + x + \{T\} = \{S'\} + \{T\} = \{S' \cup T\}, \end{aligned}$$

contrary to Theorem 12, since $S' \cup T$ is a *proper* subset of $S \cup T$. Thus the necessity of the condition is established.

---

exchange lattice and get many of our results in Sections 7, 8 from MacLane's theory of exchange lattices [1].

[25] Actually we can derive the slightly stronger result: *If* $\{A\} \supset b$ *there is an* $a \subset A$ *which may be exchanged for* b *without altering* $\{A\}$.

[26] $S \cdot T$ denotes the *intersection* or *logical product* of sets S, T.

To prove its sufficiency suppose $\{S\} \cdot \{T\} = O$, S and T independent but $S \cup T$ not independent. Then $\{(S \cup T) \dot{-} x\} \supset x$ for some $x \subset S \cup T$. For definiteness suppose $x \subset S$. Then we have

$$x \subset \{S' \cup T\} = \{S'\} + \{T\} \tag{1}$$

where $S' = S \dot{-} x$. Since S is independent we have

$$x \not\subset \{S'\}. \tag{2}$$

$\{T\} \neq O$, for otherwise (1) and (2) would be inconsistent. Thus from (1) we can get

$$x \subset \{S'\} + t = \{S' \cup t\} \tag{3}$$

where $t \subset \{T\}$. Applying Lemma 1 to (2) and (3) we have

$$\{S\} = \{S' \cup x\} = \{S' \cup t\} \supset t$$

so that $t \subset \{S\} \cdot \{T\}$, contrary to hypothesis. Thus the theorem is proved.[27]

COROLLARY 1. *If S is independent and $S \cup x$ is not independent then $\{S\} \supset x$.*[28]

Proof. S and x are independent, but $S \cup x$ is not independent. Hence $\{S\} \cdot \{x\} \neq 0$.

COROLLARY 2. *Let S be a maximal independent subset of group G. Then S is a basis of G.*[29]

Proof. Let $x \subset G$. If $x \subset S$ certainly $x \subset \{S\}$. If $x \not\subset S$ then $S \cup x$ is not independent and $x \subset \{S\}$ by Corollary 1.

7. The exchange theorem. Given a set of generators of a linear space or group we often wish to shift to a different set of generators. For example if "plane" π is determined by a_1, a_2, a_3 we can also determine π by any three of its elements b_1, b_2, b_3, which are not on a line; i. e. in $\{a_1, a_2, a_3\}$ we can exchange a_1, a_2, a_3 for b_1, b_2, b_3 getting $\{a_1, a_2, a_3\} = \{b_1, b_2, b_3\}$. This and similar situations suggest the following exchange theorem which is the key to the study of linear independence and bases of groups and, in essence, is the well known exchange theorem of Steinitz in the algebraic theory of dependence:[30]

THEOREM 14. *Let $\{a_1, \cdots, a_m\} \supset b_1, \cdots, b_n$ where $b_1, \cdots, b_n$ are*

[27] Observe that the condition $S \cdot T = O$ is not used in the sufficiency proof.
[28] Cf. van der Waerden [1], p. 95 Satz 2.
[29] Cf. van der Waerden [1], p. 96 Satz 3, especially the proof.
[30] van der Waerden [1], p. 96 Satz 4.

distinct and form an independent set. Then there exist n of the a's, $a_{i_1}, \cdots, a_{i_n}$ *which may be exchanged for* $b_1, \cdots, b_n$ *respectively without altering* $\{a_1, \cdots, a_m\}$, i. e. $\{a_1, \cdots, a_m\} = \{b_1, \cdots b_n, a_{i_{n+1}}, \cdots, a_{i_m}\}$.[31]

Proof. If $n = 0$ the theorem is trivial. Suppose $n > 0$. Our procedure is to exchange a's for b's one at a time by Lemma 2, until n a's have been exchanged. By hypothesis $\{a_1, \cdots, a_m\} \supset b_1$, so that by Lemma 2 for $p = 0$ we may exchange one of the a's for b_1 getting

$$(1) \quad \{a_1, \cdots, a_m\} = \{a_1, \cdots, a_{i-1}, b_1, a_{i+1}, \cdots, a_m\} = \{b_1, a_{j_2}, \cdots, a_{j_m}\}$$

where $a_{j_2}, \cdots, a_{j_m}$ is a permutation of $a_1, \cdots, a_{i-1}, a_{i+1}, \cdots, a_m$.

Similarly (1) and the hypothesis imply $\{b_1, a_{j_2}, \cdots, a_{j_m}\} \supset b_2$. Since b_1, b_2 are distinct by hypothesis and form an independent set by Theorem 11 we have $\{b_1\} \not\supset b_2$. Thus by Lemma 2 for $p = 1$, we may exchange one of the a_j's for b_2 getting

$$\{b_1, a_{j_2}, \cdots, a_{j_m}\} = \{b_1, b_2, a_{k_3}, \cdots, a_{k_m}\}$$

and by (1)

$$\{a_1, \cdots, a_m\} = \{b_1, b_2, a_{k_3}, \cdots, a_{k_m}\}$$

where $a_{k_3}, \cdots, a_{k_m}$ denote the "unexchanged" a's.

Repeating this process we get finally

$$\{a_1, \cdots, a_m\} = \{b_1, \cdots, b_n, a_{i_{n+1}}, \cdots, a_{i_m}\}.$$

An immediate consequence of the exchange theorem is

COROLLARY 1. *If* $\{a_1, \cdots, a_m\} \supset b_1, \cdots, b_n$ *where* $b_1, \cdots, b_n$ *are distinct and form an independent set then* $m \geqq n$.[32]

From this we can easily derive

COROLLARY 2. *If group G has a finite basis, the number of elements in a basis of G is uniquely determined.*

This result suggests

DEFINITION 8. *If group G has a finite basis we say that G is of finite rank or briefly is finite, otherwise G is of infinite rank or is infinite.*[33] *If G is finite*

[31] Cf. Menger [2], p. 463 *the fundamental theorem on independent points.*

[32] In developing the consequences of the exchange theorem through Theorem 15 we follow van der Waerden [1], pp. 96, 97 quite closely.

[33] Cf. terms *finite, infinite extension of a field.*

the number of elements in a basis of G is its rank or dimension.[34] *We denote the dimension of G functionally by* $d(G)$.

The remainder of this section is devoted to *finite* groups. We consider infinite groups in the next section. First we restate the last corollary: *A finite group has a uniquely determined rank.* We continue with further corollaries of the exchange theorem.

COROLLARY 3. *If group G has rank n, any independent set of n elements of G is a basis of G.*

COROLLARY 4. *Let A be an independent set of n elements. Then there is one and only one group of rank n containing A.*[35]

We establish a few simple properties of subgroups of a finite group in

THEOREM 15. *Let A be a subgroup of a finite group G. Then* (a) *A is a finite group*; (b) *any basis of A can be extended to form a basis of G*; (c) $d(A) \leqq d(G)$, *the equality subsisting if and only if* $A = G$.

Proof. (a) Let $G = \{a_1, \cdots, a_m\}$. By Corollary 1 of Theorem 14 an independent subset of G contains *at most* m elements. Hence we can find S, a *maximal* independent subset of A. S is a finite set and by Corollary 2 of Theorem 13 it is a basis of A. Thus the group A is finite by Definition 8.

(b) Similarly if S' is any basis of A, and T is a maximal independent subset of G which contains S', then T is a basis of G.

(c) This follows readily from $T \supset S'$.

The last theorem leads readily to the following property of *complementation,* well known in Boolean algebra:

THEOREM 16. (Complementation) *Let A be a subgroup of a finite group G. Then there exists a group X such that* $A + X = G$, $A \cdot X = O$.

Proof. By the last theorem G has a basis of the form $S_1 \cup S_2$, where S_1 is a basis of A and $S_1 \cdot S_2 = O$. Thus

$$G = \{S_1 \cup S_2\} = \{S_1\} + \{S_2\} = A + X$$

where $X = \{S_2\}$. Furthermore $S_1 \cup S_2$ is independent so that by Theorem 13, $A \cdot X = \{S_1\} \cdot \{S_2\} = O$.

---

[34] As defined, the dimension of a linear space (group) exceeds by unity the familiar geometrical notion of dimension. Thus an element (point) has dimension 1, a line has dimension 2 etc. This is very natural in the present context.

[35] Cf. Veblen and Young [1], p. 32 Corollary.

Many theorems of elementary projective geometry deal with the intersection of linear spaces contained in a given linear space, e. g. *two lines in a plane meet in a point.* We now derive a general intersection theorem for linear spaces of finite dimensions:

THEOREM 17. (Dimension Formula). *If A, B are finite subgroups of a given group*

$$d(A+B)+d(A\cdot B)=d(A)+d(B).^{36}$$

Proof. First we suppose $A\cdot B=O$. The conclusion then takes the form

$$d(A+B)=d(A)+d(B). \tag{1}$$

Let S, T be bases of A, B respectively. Then we have

$$A+B=\{S\}+\{T\}=\{S\cup T\}. \tag{2}$$

$S\cup T$ is an independent set by Theorem 13 since S, T are independent and $\{S\}\cdot\{T\}=A\cdot B=O$. Thus by (2) $S\cup T$ is a basis of $A+B$.

Observe that $S\cdot T=O$. Denoting the cardinal number of set X by $\bar{X}$ we have

$$d(A+B)=\overline{S\cup T}=\bar{S}+\bar{T}=d(A)+d(B),$$

so that (1) holds when $A\cdot B=O$.

Now we assume no restriction on $A\cdot B$. Since B is a *finite* group we have by Theorem 16

$$B=A\cdot B+X, \qquad (A\cdot B)\cdot X=O \tag{3}$$

for some subgroup X of B. Thus we may apply (1) to $A\cdot B, X$ getting

$$d(B)=d(A\cdot B+X)=d(A\cdot B)+d(X). \tag{4}$$

Moreover we have

$$A+B=A+A\cdot B+X=A+X \tag{5}$$

by the absorption principle. Since $X\subset B$ we have $X=B\cdot X$ so that $A\cdot X=A\cdot B\cdot X=O$ by (3). Thus we may apply (1) to A, X and we get in view of (5)

$$d(A+B)=d(A+X)=d(A)+d(X). \tag{6}$$

Eliminating $d(X)$ between (4) and (6) we get the desired result.

---

[36] This is essentially Theorems $S_n 2$, $S_n 3$ Veblen and Young [1], pp. 32, 33. We are following a proof of von Neumann for the same result in continuous geometry [1] Theorem 6.9 part 2.

8. Infinite groups. In this section we extend some of our results on dimensionality and complementation to groups of *infinite* rank. As might be expected we employ the axiom of choice and assume that any set can be well-ordered.

First we generalize Corollary 2 of Theorem 14.

THEOREM 18. *Any two bases of group G have the same cardinal number.*[37]

Proof. Let S, T be bases of G. In view of Corollary 2 of Theorem 14 we may assume that G is of infinite rank. Then S, T are infinite sets.

Let $s \subset S$. Then $s \subset \{S\} = \{T\}$ so that by Theorem 9, $s \subset t_1 + \cdots + t_n$ where the t's are in T. Hence, assuming Zermelo's Principle, we may associate with each s a *finite* set $T_s \subset T$ such that s is contained in the sum of the elements of T_s. Let T' denote the set theoretic sum $\mathop{\mathfrak{d}}_{x \subset S}(T_x)$. Then by Theorem 9, $s \subset \{T'\}$ so that $S \subset \{T'\}$. Hence $\{T'\} \supset \{S\} = G$ and T' is a set of generators of G. But $T' \subset T$. Thus $T' = T$ by Theorem 12, since T is independent. Hence, representing the cardinal number of X by $\bar{X}$ and cardinal summation by $\mathfrak{d}$, we have

$$\bar{T} = \bar{T}' \leqq \mathop{\mathfrak{d}}_{x \subset S}(\bar{T}_x) \leqq \aleph_0 \cdot \bar{S} = \bar{S}$$ [38]

since each T_x is finite and S is infinite. In the same way we show $\bar{S} \leqq \bar{T}$. Hence $\bar{S} = \bar{T}$ and the theorem is established.

A finite group has a basis *by definition*. We now prove

THEOREM 19. *Any group G has a basis.*[39]

Proof. Let G be well-ordered. Let S be the set of $x \subset G$ such that $x \not\subset \{G_x\}$ where G_x is the *section of G determined by x.*[40] We show that S is a basis of G.

First we show that S is a set of generators of G. Suppose this is false. Then there would be a first element $a \subset G$ such that $a \not\subset \{S\}$. Hence $a \not\subset S$ so that $a \subset \{G_a\}$ by definition of S. Moreover all elements of G which precede a are in $\{S\}$. Thus $G_a \subset \{S\}$ so that $a \subset \{G_a\} \subset \{S\}$. This contradiction establishes that S is a set of generators of G.

We merely have to show that S is independent. Suppose this false. Then $s \subset \{S \dot{-} s\}$ for some $s \subset S$. Thus by Theorems 9 and 10, $s \subset s_1 + \cdots + s_m$

---

[37] Cf. Steinitz [1], p. 124, Theorem 3, MacLane [1], Theorem 6. Our proof is simpler.

[38] See Hausdorff [1], p. 71.

[39] Cf. Steinitz [1], p. 118, Theorem 1. Essentially we follow his proof.

[40] That is the set of elements which precede x in the well-ordering of G.

$= \{s_1, \cdots, s_m\}$ where $s_i \subset S \dot{-} s$. Hence $s, s_1, \cdots, s_m$ form a finite subset of S which is *not* independent. Let $s'_1, \cdots, s'_n$ form such a subset of S containing as few elements as possible and suppose that s'_i precedes s'_j if $i < j$. Then $s'_1, \cdots, s'_{n-1}$ constitute an independent set so that by Corollary 1 of Theorem 13

$$s'_n \subset \{s'_1, \cdots, s'_{n-1}\} \subset \{G_{s'_n}\}$$

contrary to the definition of S. Thus S is independent and the theorem is established.

We extend Definition 8 to cover all groups:

DEFINITION 8′. *The rank or dimension of group G is the cardinal number of a basis of G.*

Thus we may combine and restate Theorems 18 and 19: *Any group has a uniquely determined rank.*

Now we remove the *finiteness* condition from Theorem 16 getting

THEOREM 20. (Complementation) *Let A be a subgroup of G. Then there exists a group X such that $A + X = G$, $A \cdot X = O$.*

Proof. Consider a well-ordering of G in which the elements of A precede those not in A. We define set S exactly as in Theorem 19. Let $S = S_1 \cup S_2$ where $S_1 = S \cdot A$ and $S_1 \cdot S_2 = O$. Exactly as in Theorem 19 we prove S is a basis of G, and S_1 is a basis of A. Thus, as in Theorem 16, $S_1 \cup S_2$ is a basis of G, S_1 is a basis of A and $S_1 \cdot S_2 = O$. The result now follows exactly as in Theorem 16.

Consider now the derivation of the dimension formula in Theorem 17. The hypothesis that groups A, B be finite is used only in order to apply Theorem 16. Thus in view of Theorem 20 we may assert the dimension formula for A, B of arbitrary rank:

THEOREM 21. *If A, B are subgroups of a given group*

$$d(A + B) + d(A \cdot B) = d(A) + d(B).$$

9. Boolean and projective groups. In J2 we have postulated $a + b \supset a, b$. The simplest possible type of group would be one in which $a + b$ contains only a and b. This suggests postulate

J7. $a + b = a \cup b$.

J7 is obviously equivalent to: *Each line contains exactly two points.*

Consider a group G which satisfies J7. Any subset A of G is a subgroup

of G. For $A \supset a, b$ implies $A \supset a \cup b = a + b$. Moreover, if A, B are subgroups (i. e. subsets) of G, $A + B$ is merely $A \cup B$. Thus our system is essentially the Boolean algebra of all subsets of G.[41,42] It is easily seen for any $A \subset G$, that $\{A\} = A$ and A is an independent set. Thus (set) A is a basis of (group) A and the rank of A is its cardinal number. The dimension formula thus turns out to be a very familiar property of cardinal number in set theory.[43]

Now let us substitute for J7, the postulate

J7′. *If* $a \neq b$, $\quad a + b \neq a \cup b$.

This is obviously equivalent to PG4 (Section **1**), so that by our discussion in Section **2**, J1, · · ·, J6, J7′ and PG1, · · ·, PG4 are equivalent. Thus classical projective geometries are characterized as groups satisfying J7′.

The foregoing discussion suggests

DEFINITION 9. *A group satisfying* J7 (J7′) *is called a Boolean* (*projective*) *group.*

A deeper, group theoretic, characterization of these ideas is given in the next section.

10. Homomorphisms and congruence relations. In this section we study *homomorphisms* of groups and certain related ideas.

DEFINITION 10. *Let f be a single-valued mapping of a group G on a group G' such that $f(x + y) = f(x) + f(y)$. We say that f is a homomorphism of G and that G is homomorphic to G'. If f is one-to-one we use the terms isomorphism and isomorphic in a similar way.*

Let f be a homomorphism of G and let us define $x \equiv y$ (*x is congruent to y*) to mean $f(x) = f(y)$. It is easy to see that for arbitrary elements of G: (a) $x \equiv x$; (b) $x \equiv y$ *implies* $y \equiv x$; (c) $x \equiv y$, $y \equiv z$ *imply* $x \equiv z$; (d) $x \equiv y$, $x' \equiv y'$ *imply* $x + x' \equiv y + y'$.[44] This suggests

DEFINITION 11. *Let $\equiv$ be a relation defined in group G satisfying* (a), (b), (c), (d) *above. Then we say that $\equiv$ is a congruence relation in G.*[45] *If*

---

[41] Veblen and Young [1], p. 33 in essence asserts this.

[42] Conversely, if G is any set the Boolean algebra of all subsets of G may be converted into a group satisfying J7 by defining $a + b = a \cup b$ for arbitrary elements a, b of G.

[43] Veblen and Young, *ibid.* suggests this property as an aid in remembering Theorems S_n2, S_n3 which are equivalent to the dimension formula.

[44] We interpret statements of the form $S \equiv T$ where S, T are sets to mean that *every element of each set has the relation $\equiv$ to some element of the other set.*

[45] Cf. Birkhoff [2], p. 3 where the operations are assumed to be single-valued.

$x \equiv y$ is defined by $f(x) = f(y)$ where f is a homomorphism of G, we say that $\equiv$ is the congruence relation determined by f. Any group G admits the trivial congruence relations (1) *$a \equiv b$ if $a = b$ and* (2) *$a \equiv b$ for any $a, b \subset G$.*

In virtue of (a), (b), (c) a congruence relation $\equiv$ effects a separation of G into a disjoint set of maximal classes of congruent elements which we call the *residue classes* of $\equiv$. Let $G(\equiv)$ denote this set of residue classes. In $G(\equiv)$ we define the operation $+$ thus: $R(x) + R(y) = R(x + y)$, where $R(x)$ denotes the residue class containing x and $R(x + y)$ is the *set* of residue classes which contain elements of $x + y$. In virtue of (d) $R(x) + R(y)$ is independent of the choice of the elements x, y determining $R(x)$, $R(y)$ respectively. Thus $x \to R(x)$ is a mapping of G on $G(\equiv)$ which preserves sums and it is easy to show that $G(\equiv)$ *is a group* and that G *is homomorphic to* $G(\equiv)$. Thus to $\equiv$, any congruence relation in G, is associated a homomorphism mapping G into $G(\equiv)$; moreover the congruence relation determined by this homomorphism is precisely $\equiv$. Finally if G is homomorphic to group G' and $\equiv$ is the congruence relation determined by the homomorphism, G' is isomorphic to $G(\equiv)$. Thus there is essentially an equivalence between the homomorphisms of G and its congruence relations, and it is immaterial in theory which we study. We find it more convenient to study congruence relations.[46]

In classic group theory a subgroup is *normal* if and only if it determines a congruence relation (or equivalently a homomorphism). This suggests

DEFINITION 12. *If A is a subgroup of G and $x + A = y + A$ we write $x \equiv y \pmod{A}$. If this relation "congruence modulo A" is a congruence relation in G we call it modular and say that A is a normal subgroup of G. Any group G has the trivial normal subgroups G and O which give rise to the trivial congruence relations mentioned in the last definition. Group G is simple if it has only trivial normal subgroups.*

We proceed to study *normal* subgroups. Consider

$$x \equiv y \pmod{A} \tag{1}$$

where A is a normal subgroup of G. (1) is obviously satisfied if (a) $x, y \subset A$ or (b) $x = y$. Suppose (1) true but (a) and (b) false. Then (1) implies $y \subset x + A$ in which $A \neq O$. Thus $y \subset x + a$ where $a \subset A$ so that $a \subset x + y$. Since A is normal, (1) implies $x + y \equiv x + x \pmod{A}$ so that $a \equiv x \pmod{A}$. Hence $A = a + A = x + A \supset x$ contrary to our supposition. Thus (1) implies (a) or (b) and we may assert

[46] The relations between homomorphisms and congruence relations sketched above hold in a very general type of abstract algebra. See Birkhoff [2], pp. 2, 3.

THEOREM 22. *Let A be a normal subgroup of G. Then $x \equiv y \pmod{A}$ is equivalent to the disjunction $x, y \subset A$ or $x = y$.*

Now consider $a + b$ where $a \subset A$, $b \not\subset A$. Let $x \subset a + b$, $x \neq a$. Then $a + x = a + b$ so that $A + x = A + b$ and $x \equiv b \pmod{A}$. Thus $x = b$ by Theorem 22, and $a + b = a \cup b$.

This suggests

THEOREM 23. *A subgroup A of G is normal if and only if $a + b = a \cup b$ for $a \subset A$, $b \not\subset A$.*[47,48]

Proof. We have just established the necessity of the condition. To prove the sufficiency suppose

$$a + b = a \cup b \quad \text{if} \quad a \subset A, b \not\subset A. \tag{1}$$

We show that A is normal. Properties (a), (b), (c) of Definition 11 obviously are satisfied by the relation $x \equiv y \pmod{A}$. Before considering property (d) we show that $x \equiv y \pmod{A}$ implies $x, y \subset A$ or $x = y$. To prove this we suppose $x \equiv y \pmod{A}$, $x \neq y$ and show $x, y \subset A$. We have $x + A = y + A$ where $A \neq O$. Thus $x \subset y + A$. If $y \subset A$ then $y + A = A$ and $x \subset A$. Suppose $y \not\subset A$. We have $x \subset y + a$ for some $a \subset A$. By (1) $y + a = y \cup a$, so that $x = a \subset A$. By symmetry $y \subset A$.

Now to prove (d) suppose $x \equiv y \pmod{A}$, $x' \equiv y' \pmod{A}$. Then as just shown $x, y \subset A$ or $x = y$; likewise $x', y' \subset A$ or $x' = y'$. If $x, y, x', y' \subset A$ then $x + x'$, $y + y' \subset A$ and we have

$$x + x' \equiv y + y' \pmod{A} \tag{2}$$

since all elements of A are congruent modulo A. Suppose that only one of the statements $x, y \subset A$ and $x', y' \subset A$ holds. Let us say the former. Then $x', y' \not\subset A$ and $x + x' = x \cup x'$, $y + y' = y \cup y'$ by (1) so that (2) holds. Finally if both $x, y \subset A$ and $x', y' \subset A$ are false then $x = y$, $x' = y'$ and (2) is trivial. Thus A is a normal subgroup of G by definition and the proof is complete.

COROLLARY 1. *G is a Boolean group if and only if all its subgroups are normal.*

COROLLARY 2. *A projective group is simple.*

---

[47] Thus the theorem of classic group theory, that any subgroup of an abelian group is normal, does not hold here.

[48] In geometrical language, *every line joining a point in A to a point not in A consists of two points.*

COROLLARY 3. *If A is a normal subgroup of G we can express G as $A + B$ where $A \cdot B = O$ and B also is a normal subgroup of G.*[49]

Proof. Taking B as the *set complement* $G \dot{-} A$ we have $a + b = a \cup b$ for $a \subset A$, $b \subset B$. Hence $A + B = A \cup B = G$ and $A \cdot B = O$. In view of Theorem 23, B is normal if it is a subgroup of G. To show the latter suppose $B \supset x, y$ and $x + y \supset z$ but $B \not\supset z$. Then x, y, z are distinct and $A \supset z$. Hence $y \subset x + y = x + z = x \cup z$ so that $y = x$ or z. This contradiction establishes the result.

In classical group theory *all* congruence relations (or homomorphisms) are determined by normal subgroups. This is not true here. However we have

THEOREM 24. *Any congruence relation in group G is expressible as a disjunction of modular congruence relations.*

Proof. Let $\equiv$ be a congruence relation in G and R any residue class of $\equiv$. First we show that R is a subgroup of G. Let $x, y \subset R$, $z \subset x + y$. Then $x \equiv y$ and since $x \equiv x$ we have $x = x + x \equiv x + y \supset z$. Hence $z \equiv x$ and so $z \subset R$. Thus R is closed under $+$ and is a subgroup of G.

Now supposing $d(R) > 1$, we show that R is normal. Let $a \subset R$, $b \not\subset R$. By Theorem 23 we need show merely $a + b = a \cup b$. Suppose $a + b \neq a \cup b$. Then $a + b \supset c$ for some $c \neq a, b$ so that $a + b = a + c$. Let $a' \subset R$, $a' \neq a$. Then $a \equiv a'$ so that $a + b \equiv a' + b$. Thus $c \subset a + b$ implies $c \equiv c'$ where $c' \subset a' + b$. $c \neq c'$, for otherwise using J6, $a + b = b + c = b + c' = a' + b \supset a'$ so that $b \subset a + b = a + a' \subset R$. Likewise $c' \neq a'$ for otherwise $c \equiv c' = a' \equiv a$ so that $R \supset c$ and $R \supset a + c = a + b \supset b$. Hence $a' + b = a' + c'$. By the dimension formula

$$d[(c + c') \cdot (a + a')] = d(c + c') + d(a + a') - d(c + c' + a + a').$$

Moreover

$$\begin{aligned} d(c + c' + a + a') &= d[(a + c) + (a' + c')] \\ &= d[(a + b) + (a' + b)] = d(a + a' + b) = 3 \end{aligned}$$

since a, a', b are distinct and form an independent set. Thus $d[(c + c') \cdot (a + a')] = 2 + 2 - 3 = 1$ and $(c + c') \cdot (a + a') \neq O$. Let R' be the residue class of $\equiv$ which contains c and c'. Then $R' \cdot R \supset (c + c') \cdot (a + a') \neq O$ so that $R' = R$. Thus $R \supset a + c \supset b$. This contradiction implies that R is a normal subgroup of G.

Let α be the set of residue classes R for which $d(R) > 1$. If α is empty,

---

[49] This suggests the idea, *direct product of two classic groups*. We could define G as the *direct sum* of its subgroups A, B if A, B are normal, $G = A + B$, $A \cdot B = O$. Cf. van der Waerden [1], p. 141.

$\equiv$ is the relation *identity* and the result is trivial. Suppose that α is not empty. Consider the relations $a \equiv b \pmod{R}$ and their *disjunction*, viz. $a \equiv b \pmod{R}$ *for at least one* R *in* α. If the latter relation holds $a = b$ or $a, b \subset R$ for some R in α by Theorem 22, and certainly $a \equiv b$. Conversely $a \equiv b$ implies $a = b$ or $a, b \subset R$ for some R in α, so that $a \equiv b \pmod{R}$ for some R in α. Thus $\equiv$ may be determined as the disjunction of the relations $a \equiv b \pmod{R}$ and the theorem is proved.

Actually we have shown that *any congruence relation in* G *is a disjunction of modular congruence relations determined by disjoint subgroups of* G. Conversely it can be shown that *any such disjunction is a congruence relation.* Moreover *any congruence relation in* G, *other than the identity relation, is uniquely expressible as such a disjunction if we exclude the relation identity from the terms of the disjunction.*

We now introduce an idea used by Birkhoff in his lattice formulation of finite-dimensional projective geometry, which is important in the remainder of this section.

DEFINITION 13. *Let group* $G \supset x, y$. *If* $x = y$ *or* $x + y \neq x \cup y$ *we say that* x *is perspective to* y *and we write* $x \sim y$.[50]

THEOREM 25. *Perspectivity is a congruence relation in group* G.[51]

Proof. Properties (a), (b) of Definition 11 obviously hold. Suppose $x \sim y$, $y \sim z$. We show $x \sim z$. We assume x, y, z distinct and that they form an independent set, since the other cases are trivial. We have $x + y \supset p \neq x, y$ and $y + z \supset q \neq y, z$. We use the dimension formula, as in Theorem 24 to show $(x + z) \cdot (p + q) \neq O$. Thus there exists $r \subset x + z$, $p + q$. If $r = x$ or z we can show by J6 that x, y, z are not independent. Thus $r \neq x, z$ and $x \sim z$ by definition.

We have to show only (d) of Definition 11. Suppose $x \sim y$, $x' \sim y'$. Every element of $x + x'$ is perspective to x or to x'. Similarly every element of $y + y'$ is perspective to y or y'. Hence

$$x + x' \sim x \cup x' \sim y \cup y' \sim y + y',$$

and our theorem is established.

Consider the residue classes into which $\sim$ separates G. We know from our discussion in Theorem 24 that the residue classes are subgroups of G. If x, y are in distinct residue classes, $x \sim y$ is false so that $x + y = x \cup y$. From

---

[50] Birkhoff [1], p. 746 calls the relation *conjointness*. Cf. Birkhoff [2], Definition 4.2 and Lemma p. 61. Observe that $x \sim y$ if and only if x, y belong to a common *projective* subgroup of G.

[51] Cf. Birkhoff [1], Lemma 2.

this it readily follows that $G(\sim)$ *is a Boolean group.*[52] Moreover by Theorem 23 the residue classes are normal subgroups of G. Further, if $x \neq y$ are in the same residue class then $x + y \neq x \cup y$ so that the residue classes are *projective* and hence *simple* subgroups of G, by Corollary 2 of Theorem 23. Thus we may state

COROLLARY 1. *Any group G may be expressed as a set union of disjoint normal subgroups which are simple.*[53,54]

Now we can easily characterize *projective groups* and thus *projective geometries.*

COROLLARY 2. *A group is projective if and only if it is simple.*[55]

Proof. We have already shown the necessity of the condition (Corollary 2, Theorem 23). Suppose that the group G is simple. Let G be decomposed by the relation $\sim$ as in Corollary 1. Then the decomposition involves only one term which as we saw must be a projective group. Thus G is projective.

Now consider *any* decomposition of the group G, $G \neq O$, into disjoint simple normal subgroups $\neq O$. By Corollary 2 these subgroups are projective. Hence x, y are in the same term of the decomposition if and only if $x \sim y$. Thus the terms in the decomposition are the residue classes of $\sim$ and we may assert

COROLLARY 3. *The decomposition in Corollary* 1 *is unique if G and the terms into which it is decomposed are not empty.*

11. The inverse operation. Thus far we have focused our attention on the operation $+$. We have not employed the inverse operation (although it is sometimes convenient to do so) in order to stay close to the primitive geometrical notion *join.* Now in order to compare our idea of group with that of classical abelian group we introduce

DEFINITION 14. *If* $a, b \subset G$ *and* $a \neq b$[56] *the set of* x *satisfying* $b + x \supset a$ *is called* $a - b$.

---

[52] It can be shown that $\sim$ is the "weakest" of the relations $\equiv$ for which $G(\equiv)$ is Boolean, i. e. $x \sim y$ implies $x \equiv y$ for any such $\equiv$.

[53] The use of the relation $\sim$ to effect this decomposition is suggested by Birkhoff's derivation of a similar theorem, [1] p. 747, Theorem 3. Cf. Menger [2], Theorem, p. 473.

[54] If the definition of direct sum suggested in the footnote to Corollary 3 of Theorem 23 is extended in a natural way to an arbitrary set of terms, we can restate this result in the form: *Any group is a direct sum of simple groups.* Cf. Birkhoff [2], p. 60, Corollary.

[55] Cf. Birkhoff's characterization of finite-dimensional projective geometries [2], Theorem 4.12.

[56] If we allow $a = b$ we get by J2, $a - a = G$. Thus $a - a$ is comparable to $0 \div 0$ in a field and so is not defined.

Suppose $a \neq b$. We observe by J2 that $a - b \supset a$, so that $a - b \neq 0$. Further if $x \subset a - b$ we have $b + x \supset a$ and by J6, $a + b = b + x \supset x$ so that $a + b \supset a - b$. Thus we may assert

THEOREM 26. *If $A \subset G$ is closed under $+$, A is also closed under $-$.*[57]

Thus our definition of subgroup is essentially the familiar one.[58] Furthermore Theorem 26, which is derived using J6, can be shown equivalent to J6 in the presence of J1, $\cdots$, J5. Thus the role of J6 becomes more apparent. J6 or its equivalent, Theorem 26, characterizes our type of group as one of the simplest systems satisfying J1, $\cdots$, J5, in a sense analogous to classical finite abelian groups which also satisfy Theorem 26.

BROOKLYN COLLEGE, BROOKLYN, N. Y.

REFERENCES

G. Birkhoff:
[1] "Combinatorial relations in projective geometries," *Annals of Mathematics*, vol. 36 (1935), pp. 743-748.
[2] *Lattice Theory*, New York, 1940.

F. Hausdorff:
[1] *Mengenlehre*, second edition, Berlin, 1927.

S. MacLane:
[1] "A lattice formulation for transcendence degrees and p-bases," *Duke Mathematical Journal*, vol 4 (1938), pp. 455-468.

K. Menger:
[1] "Bemerkungen zu Grundlagenfragen," IV, *Jahr. D. M.-V.*, vol. 37 (1928), pp. 309-325.
[2] "New foundations of projective and affine geometry," *Annals of Mathematics*, vol. 37 (1936), pp. 456-82.

J. von Neumann:
[1] *Continuous Geometry*, part I, mimeographed, Princeton, 1936.

E. Steinitz:
[1] *Algebraische Theorie der Körper*, Berlin, 1930.

O. Veblen and J. W. Young:
[1] *Projective Geometry*, vol. 1, Boston, 1910.

B. L. van der Waerden:
[1] *Moderne Algebra*, vol. 1, first edition, Berlin, 1930.

[57] That is $A \supset x, y$ implies $A \supset x - y$. Observe that if $a - a$ had been defined the only subsets of G closed under $-$ would be O, G.

[58] We can make this a little more precise. By Theorem 26 our definition of subgroup is equivalent to: *$A \subset G$ is a subgroup of G if it is closed under addition of any two elements and subtraction of distinct elements.* This property is easily seen to characterize subgroups of classical abelian groups (expressed additively).

SYSTEMS OF RATIONAL CURVES.*

By JULIAN L. COOLIDGE.

The following investigation was prompted by this theorem:

BLUTEL'S THEOREM 1. *If the conics of an analytic one-parameter system be not coplanar, but touch two curves, the curves conjugate to them on the surface generated will establish a projective correspondence among them.*[1]

It is evident that this theorem comes under some general classification of theorems about rational curves. Given a set of rational curves, what can be said about another set of curves that establish a projective correspondence among them? It is the purpose of the present paper to study this question.

Suppose that in three-space we have a one-parameter analytic family of rational curves. If a point have homogeneous tetrahedral coördinates x^i we may express our curves by the equations

$$(1) \qquad x^i = b_j{}^i (t)^j; \qquad b_j{}^i = b_j{}^i(v), \qquad i = 1 \cdots 4, \quad j = 0 \cdots n.$$

The parenthesis around the letter t is to indicate that the j superior is an actual exponent, not a mere index like the j inferior, or the i superior. The coefficients are supposed to be analytic in v and the ratios of the four x's are not all constants.

The differential equation of the curves conjugate to the rational ones on the surface generated may be written in a slight modification of the classical Gauss notation

$$(2) \qquad Ddt^2 + 2D'dtdv + D''dv^2 = 0.$$

(3)

$$D = | \, x x_t x_v x_{tt} \, | \text{, order } 4n - 6;$$
$$D' = | \, x x_t x_v x_{tv} \, | \text{, order } 4n - 4;$$
$$D'' = | \, x x_t x_v x_{vv} \, | \text{, order } 4n - 2.$$

The essential thing to notice here is that the orders in t of these three polynomials form an arithmetic progression of constant difference 2.

* Received October 4, 1941.

[1] Blutel, "Recherches sur les surfaces qui sont en même temps lieux de coniques et enveloppes de cônes," *Annales de l'École Normale Supérieure* (3) Tome 7 (1890), p. 155. See also Coolidge, *Transactions of the American Mathematical Society*, vol. 48 (1940), p. 365.

The differential equation of the curves conjugate to the rational ones is

$$Ddt + D'dv = 0. \tag{4}$$

Now if D' be divisible by D this will be a Riccati equation, and in consequence, by a classical theorem, the cross ratio of any four solutions is independent of v. But what does it mean that D' should be divisible by D? The vanishing of D means either that the point in question is a cusp $x^i = \rho x_t{}^i$ or an inflection $x^i = px_t{}^i + qx_{tt}{}^i$ or that the osculating plane is tangent to the surface generated. On the other hand when $D = D' = 0$ the asymptotic directions fall together, which means in metrical terms that the total curvature is 0.

THEOREM 2. *If a one-parameter family of curves have the property that the inflections, and points where the osculating plane is tangent to the surface, are all points where the asymptotic directions fall together, then these curves will be projectively related by their conjugates on the surface generated.*

It is worth attempting to find a canonical form for such sets of curves, by a suitable change of parameter. Suppose that t is an analytic function of u and v so that u is analytic in t and v. We find by straight substitution

$$| xx_u x_v x_{uv} | = \left(\frac{\partial t}{\partial u}\right)^2 \left[\frac{\partial t}{\partial v} | xx_t x_v x_{tt} | + | xx_t x_v x_{tv} |\right].$$

Put
$$t = \frac{p(v)u + q(v)}{r(v)u + s(v)} \qquad ps - qr \not\equiv 0;$$

then ρx^i will be a polynomial of order n in u and the cross ratio of four values for t is that of the corresponding values for u. But when the conditions for Theorem 2 are fulfilled

$$| xx_t x_v x_{tv} | \equiv (A(v)t^2 + B(v)t + C(v)) | xx_t x_v x_{tt} |$$

$$= \frac{\alpha u^2 + \beta u + \gamma}{(ru + s)^2} | xx_t x_v x_{tt} |.$$

Now

$$\frac{\partial t}{\partial v} = \frac{(rp' - r'p)u^2 + [rq' - r'q + sp' - s'p]u + sq' - s'q}{(ru + s)^2};$$

if, then, we give to p, q, r, s such values that

$$(rp' - r'p) = -\alpha \qquad rq' - r'q + sp' - s'p = -\beta \qquad sq' - s'q = -\gamma$$

$$x^i = a_j{}^i(u)^j \qquad a_j{}^i = a_j{}^i(v) \qquad j = 1 \cdots 4, \quad i = 0 \cdots n \tag{5}$$

$$| xx_u x_v x_{uv} | = 0. \tag{6}$$

It is interesting to look at the special case of Theorem 2 when the curves are plane. The coördinates of the plane will be

$$\xi_i = \rho \mid x^j x_u{}^k x_{uu}{}^l \mid .$$

When $D = 0$ we either have a cusp or an inflection, or

$$x_v{}^i = lx^i + mx_u{}^i + nx_{uu}{}^i.$$

But then

$$\xi_i x_v{}^i = 0, \qquad \xi_{iv} x^i = 0,$$

and we are at one of the n points where the curve meets the characteristic line in the plane. The cusps will count doubly in this reckoning, for we find by Plücker's equations, that if a plane curve of genus 0, a rational curve, of order n, have κ cusps, and i inflections

$$n + i = 4n - 6 - 2\kappa.$$

Now consider one of the points where the curve meets the characteristic line in its plane. Let $x^i(v)$ be its coördinates; $x_u{}^i$ will be the coördinates of some point on the tangent. If (x), (x_u), (x_v) are linearly independent, then, since $D' = 0$,

$$x_{uv}{}^i = \lambda x^i + \mu x_u{}^i + \nu x_v{}^i \qquad \xi_i x_{uv}{}^i = 0$$

But since $\xi_i x_u{}^i = 0$, $\xi_i x_{uv}{}^i + \xi_{iv} x_u{}^i = 0$ whereas $\xi_{iv} x_u{}^i \neq 0$, for the tangent crosses the characteristic line only at (x). Hence (x), (x_u), (x_v) are linearly dependent, or the moving curve touches n curves at the points where it cuts the characteristic line.

THEOREM 3. *A one parameter family of rational plane curves of order n will be projectively related by their conjugates on the surface generated if and only if their inflections are at points where the asymptotic directions coincide, and they are tangent to n curves, or if it is a limiting form of such a system.*

It is to be noted that Theorem 1 is a special case of this.

We can find a beautiful set of rational curves whose conjugates also are rational, and fulfill the conditions for Theorem 2 as follows. Let ϕ^i be a polynomial of order n in u with constant coefficients, ψ^i a polynomial of order m in v with constant coefficients, while α_i, β_k are also constants:

$$x^i = \alpha_j\psi^j\phi^i + \beta_k\phi^k\psi^i; \quad x_u{}^i = \alpha_j\psi^j\phi_u{}^i + \beta_k\phi_u{}^k\psi^i; \quad x_v{}^i = \alpha_j\psi_v{}^j\phi^i + \beta_k\phi^k\psi_v{}^i;$$
$$x_{uv}{}^i = \alpha_j\psi_v{}^j\phi_u{}^i + \beta_k\phi_u{}^k\psi_v{}^i.$$

$$\alpha_j\psi_v{}^j[\beta_k\phi_u{}^k x^i - \beta_k\phi^k x_u{}^i] - \alpha_i\psi^j[\beta_k\phi_u{}^k x_v{}^i - \beta_k\phi^k x_{uu}{}^i] = 0; \qquad D' = 0.$$

We thus have two sets of rational curves which are conjugate to one another, and cut one another projectively. Such a surface is in the nature of a generalization of a translation surface. For each value of v the corresponding curve ϕ^i will be transformed into another rational curve, each point moving along the line which connects it with the corresponding point ψ^i, with a corresponding transformation of the curves ψ^i.

Another very simple case comes when our rational curves lie on a developable surface

$$D'^2 - DD^\kappa \equiv 0.$$

Here D' is certainly divisible by D. On any developable surface, not a plane, the curves conjugate to any set are the generators.

THEOREM 4. *If a one-parameter family of rational curves lie on a non-planar developable surface, they are projectively related by the generator s.*

When the developable is a cone, the theorem is trivial; when the surface is algebraic the theorem is algebraically evident, for a one to one algebraic transformation between rational curves is projective. I have not seen an example of a non-algebraic developable surface which has a one parameter family of rational curves, but incline to the belief that such a surface exists. Here is a simple example of an algebraic surface of the sort. I begin with a cubic space curve, which is surely rational. The tangents generate a developable surface whose order is easily found to be 4. We can surely pass a one-parameter family of cubic surfaces through our curve. Each tangent to the curve will meet each surface once again, giving an algebraic residual intersection which is algebraic, in one to one algebraic correspondence with a rational curve. A plane will cut the developable in a quartic curve with three cusps, and the cubic in a cubic curve through the three cusps. There are, thus, six remaining intersections of these plane curves, so that the rational residual curves are of order 6.

We can pass from one set of rational curves fulfilling the conditions of Theorem 2 to another by a classical transformation first discovered by Laplace.[2] We start with equations (4) and replace (5) by

---

[2] Oeuvres. vol. 9. The best exposition is Tzitzeica, *Géométrie différentielle projective des reseaux*, Bucarest and Paris, 1924.

$$x_{uv}^i + ax_u^i + bx_v^i + cx^i = 0 \tag{7}$$

where a, b, c are analytic in v. We write

$$x_1^i = x_u^i + bx^i \qquad x_{1v}^i = -ax_u^i + (b_v - c)x^i.$$

This shows that as (x) moves along the curve conjugate to the rational one, (x_1) moves along the tangent to the rational one. The tangents to our rational curves will form a congruence, the focal points being (x) and (x_1). If we put

$$ab + b_v - c = h$$

we have

$$x_{iuv}^i + ax_{iu}^i + (b - (h_u/h))x_{iv}^i + (c + a_u + b_v + (ah_u/h))x^i = 0. \tag{9}$$

This shows that on the other focal surface the points (x_1) generate a set of rational curves projectively cut by their conjugates. There is a second Laplace transformation given by

$$x_2^i = x_v^i + ax^i.$$

There is a considerable literature dealing with these transformations. Our interest lies in the fact that the special property of being cut projectively by the conjugates is preserved.

In conclusion let us look for a moment at two-parameter systems of rational curves, depending on parameters v, w. Here (6) is replaced by

$$\begin{vmatrix} x & x_u & x_v + x_w(dw/dv) & x_{uv} + x_{uw}(dw/dv) \end{vmatrix} = 0,$$

$$\begin{vmatrix} x & x_u & x_v & x_{wv} \end{vmatrix} dv^2 + \left[\begin{vmatrix} x & x_u & x_v & x_{uw} \end{vmatrix} + \begin{vmatrix} x & x_u & x_w & x_{uv} \end{vmatrix} \right] dvdu + \begin{vmatrix} x & x_u & x_w & x_{uv} \end{vmatrix} dw^2 = 0. \tag{10}$$

We see that our curves can be assembled usually, in two different ways, so that they are cut projectively by their conjugates. It is interesting to inquire as to when they become identical, that is, when will the roots of (10) be equal. The Plücker coördinates of a tangent will be

$$p^{ij} = \begin{vmatrix} x^i & x^j \\ x_u^i & x_u^j \end{vmatrix}.$$

If we use the Grassmann notation

$$(p \mid q) = \Sigma\, p^{ij} q^{kl},$$

our differential equation (10) is

$$(11) \qquad (p_v \mid p_w)dv^2 + 2(p_v \mid p_w)dvdw + (p_w \mid p_w)dw^2 = 0.$$

But when u is constant, this is the differential equation for the developable surfaces of the congruence of tangents, and it is well known that a congruence of lines can be assembled into developable surfaces in only one way when, and only when, they are tangents to the asymptotic lines of some surface. This gives a curious result:

THEOREM 5. *The rational curves of a two-parameter family can be assembled in two different ways into surfaces where they are cut projectively by their conjugates, unless their tangents are the tangents to the asymptotic lines of a one parameter family of surfaces, or a limiting case of such a set.*

HARVARD UNIVERSITY.

ON THE SPECTRAL ANALYSIS OF A CERTAIN TRANSFORMATION.*

By J. L. Doob and R. A. Leibler.

1. Introduction. The analysis of a one to one measure preserving transformation [1] on an abstract space by means of the spectrum of a corresponding unitary operator goes back to Koopman.[2] His method can be described as follows:

Consider a space Ω of points ω on which is defined a Borel field [3] F_ω of sets, with $\Omega \in F_\omega$. Suppose also that there is a completely additive non-negative measure function P defined on sets of F_ω and that $P(\Omega) = 1$. Then the class $L_2(\Omega)$ of all complex valued, P-measurable, integrable squared functions of ω forms a Hermitian space.[4] If further T is a one to one measure preserving transformation of Ω into itself, the operator U defined by

$$Uf(\omega) = f(T\omega) \tag{1.1}$$

for $f \in L_\lambda(\Omega)$, is unitary and is called the unitary operator arising from T. Then by the spectral theory of unitary operators there exists a spectral family [5] $E(\lambda)$, $-\infty < \lambda < \infty$, such that: $E(\lambda) = 0$ for $\lambda \leq 0$; $E(\lambda) = 1$ (= identity) for $\lambda \geq 1$; $E(\lambda)$ is a projection operator; $E(\lambda)$ is continuous on the right except at $\lambda = 0$ and continuous on the left at $\lambda = 1$; $E(\lambda)E(\mu) = E(\mu)E(\lambda) = E(\mu)$ for $\lambda \geq \mu$; $E(0+) = \lim E(\epsilon)$ as ϵ decreases to zero is the projection on the manifold of functions invariant under U; for f and $g \in L_\lambda(\Omega)$,

* Received Sept. 2, 1941; Revised August 22, 1942.

[1] A transformation T is measure preserving if for every measurable set E, TE and $T^{-1}E$ are measurable and measure of E equals measure of TE equals measure of $T^{-1}E$.

[2] *Proceedings of the National Academy of Sciences*, vol. 17 (1931), p. 315. Koopman does not assume, as we do, that his space has finite measure.

[3] A field is a collection of sets containing with sets E and F their sum $E + F$ and their differences. A Borel field is a field containing with sets $E_1, E_2, \cdots$ their sum $E_1 + E_2 + \cdots$

[4] A Hermitian space is a space satisfying all of the postulates of a Hilbert space except those relating to separability and dimensionality. In $L_2(\Omega)$ we define the inner product of f and g by $(f, g) = \int f(\omega) \cdot \overline{g(\omega)}\, dP$ where the integral is taken over all of Ω.

[5] Cf. F. Rellich: "Spektraltheorie in nichtseparable Räumen," *Mathematische Annalen* 110 (1935), pp. 342-356. This resolution of the identity holds even if U does not arise from a measure preserving transformation.

$$(1.2) \qquad (U^n f, g) = \int_0^1 e^{2\pi i n\lambda} d(E(\lambda)f, g).$$

(Here U^n is the n-th iterate of U.) The family $E(\lambda)$ is uniquely determined by the above properties.

Koopman's idea was to characterize T in terms of the properties of $E(\lambda)$. The purpose of this paper is to make a contribution in that direction. In particular, the case where Ω is the infinite dimensional unit cube, T is the coordinate shift transformation, and P is the natural extension of n-dimensional Lebesgue measure is studied. This situation is of importance in the probability investigations of repeated trials of an experiment where each trial is independent of the others and results in a number between zero and one, and where the probability that the result is a number in the measurable set A (contained in the unit interval) is the Lebesgue measure of A.

It is shown that $E(\lambda)$ for this shift transformation is absolutely continuous after a jump at $\lambda = 0$ has been removed.[6].

The case of an abstract space Ω and a general unitary operator U in $L_2(\Omega)$ is then studied and theorems are obtained relating absolute continuity of $E(\lambda)$ with independence for the functions $U^n f$, $n = 0, \pm 1, \pm 2, \cdots$, where f is some suitably chosen element in $L_2(\Omega)$.

2. The infinite dimensional cube. Let Ω^* be the space consisting of the points ω: $(\cdots, x_{-1}, x_0; x_1, \cdots, x_j, \cdots)$ where the x_j are real numbers in the unit interval. We shall indicate how a Borel field F_ω with a measure function P can be defined on Ω^* and we shall define a transformation T which is P-measure preserving.

Consider sets E in Ω^* determined by conditions of the form

$$(2.1) \qquad x_j \varepsilon E_j, \qquad j = 0, \pm 1, \cdots, \pm n,$$

where E_j is a Lebesgue measurable set in the unit interval and n is any non-negative integer. Such sets E will play a role in Ω^* analogous to that played by intervals on the lines, rectangles (with sides parallel to the coordinate axes) in the plane, etc. Denoting the Lebesgue measure of E_j by $m(E_j)$ we define $P(E)$ by

$$(2.2) \qquad P(E) = m(E_0) \cdot m(E_1) \cdot m(E_{-1}) \cdots m(E_n) \cdot m(E_{-n}).$$

Let F denote the Borel field determined by all sets of the form (2.1). It is well known that P can be extended so as to be defined and completely additive

[6] A function of λ is absolutely continuous after a jump at $\lambda = 0$ has been removed if it is of the form $c(\lambda) + s(\lambda)$ where $c(\lambda)$ is absolutely continuous and $s(\lambda)$ is a step function continuous except at $\lambda = 0$.

on F.[7] It is then the natural extension of Borel measure in n-dimensional Euclidean space. We can now extend F to contain all subsets of sets of P-measure zero and we shall call the resulting field F_ω. If we extend P so that it is zero on subsets of sets of measure zero we have the natural extension of Lebesgue measure in n-dimensional Euclidean space.

Let T be the transformation on Ω^* defined by

$$T(\omega) = T(\cdots, x_{-1}, x_0; x_1, \cdots, x_j, \cdots) = (\cdots, x_0, x_1; x_2, \cdots, x_{j+1}, \cdots). \tag{2.3}$$

It is obvious that this transformation preserves P-measure. The transformation T is the "coördinate shift transformation" referred to in the introduction.

Let U be the unitary operator in $L_2(\Omega^*)$ arising from T. We shall show that the corresponding spectral family $E(\lambda)$ is absolutely continuous after a jump at $\lambda = 0$ has been removed. In doing this we shall make use of the following known result[8] which we state without proof:

LEMMA 1. *If U is a unitary operator in $L_2(\Omega^*)$ for which the corresponding $E(\lambda)f$ is continuous at $\lambda = \mu$, $\mu \neq 0$, then*

$$E(\mu)f - \tfrac{1}{2}E(0+)f \sim \sum_m a_m(\mu)U^m f,^9 \tag{2.4}$$

where the sign "$\sim$" denotes the fact that the left side of (2.4) *is the limit in the mean* (l. i. m.) *of the right side and where*

$$a_m(\mu) = \int_0^1 \phi_\mu(t)e^{-2\pi imt}dt = \int_0^\mu e^{-2\pi imt}dt, \tag{2.5}$$

$$\phi_\mu(t) = \tfrac{1}{2} \text{ for } t = 0 = \mu; \quad = 1 \text{ for } 0 < t < \mu; \quad = 0 \text{ for } \mu < t < 1. \tag{2.6}$$

In the space Ω^* each point ω_0 is a sequence $(\cdots, x^0_{-1}, x^0_0, x^0_1, \cdots, x^0_j, \cdots)$ of real numbers. Consider the function of ω which is equal to the j-th coördinate of ω and designate it by $x_j(\omega)$:

$$x_j(\omega_0) = x^0_j. \tag{2.7}$$

This function is real valued and its domain and range are Ω^* and the unit interval respectively. Further it is measurable and belongs to $L_2(\Omega^*)$ and

$$Ux_j(\omega) = x_{j+1}(\omega). \tag{2.8}$$

---

[7] For a proof see A. Kolmogoroff, *Grundbegriffe der Wahrscheinlichkeitsrechnung* or E. Hopf, *Ergodentheorie*.

[8] Cf. M. H. Stone, *Linear Transformations in Hilbert Space* (1932), p. 309.

[9] $\sum_j$ will stand for summation where the indicated index ranges over all integers from $-\infty$ to ∞.

The set consisting of

$$1 \text{ and } e^{2\pi i[p_0 x_j(\omega)+\cdots+p_k x_{j+k}(\omega)]}, \tag{2.9}$$

($j, p_0, \cdots, p_k$ ranging over all the integers; $p_0 \cdot p_k \neq 0$, $k = 0, 1, 2, \cdots$) is a complete orthonormal (complete o. n.) set in $L_2(\Omega^*)$. This follows from the fact that the functions of (2.9) span any manifold of $L_2(\Omega^*)$ which is determined by a finite number of the $x_j(\omega)$ and the fact that measure was defined on Ω^* so that the functions depending on only a finite number of the $x_j(\omega)$ are dense in $L_2(\Omega^*)$.

Consider the function $\phi(\omega) = e^{2\pi i[p_0 x_0(\omega)+\cdots+p_k x_k(\omega)]}$, with $p_0 \cdot p_k \neq 0$. In the following discussion, a set of integers $p_0, \cdots, p_k$, with $p_0 \cdot p_k \neq 0$ will be denoted by the letter α. The set of functions $\{U^n\phi(\omega)\}$, $n = 0, \pm 1, \pm 2, \cdots$, determines a closed linear manifold which we shall designate by $\mathfrak{M}_\alpha$, $\alpha = (p_0, \cdots, p_k)$. As α varies, we have a collection of closed linear manifolds which determine the orthogonal complement in $L_2(\Omega^*)$ of the constant functions. If we include the manifold $\mathfrak{M}_0$ of constants, we have decomposed $L_2(\Omega^*)$ into a complete set of mutually orthogonal closed linear manifolds, invariant under U.

Let f be an arbitrary function in $L_2(\Omega^*)$. Since (2.9) is a complete o. n. set in $L_2(\Omega^*)$, f has a Fourier series with respect to it. Project f on $\mathfrak{M}_\alpha$, that is, consider only the terms in the Fourier series of f which belong to $\mathfrak{M}_\alpha$. The sum of the squares of their absolute values converges by Bessel's inequality, and so by the Riesz-Fischer theorem, there exists a function $f_\alpha(t)$ in $L_2(0,1)$ [10] such that the coefficient of $e^{2\pi i[p_0 x_j(\omega)+\cdots+p_k x_{j+k}(\omega)]}$ in the Fourier series of f is

$$\int_0^1 e^{-2\pi i j t} f_\alpha(t). \tag{2.10}$$

This determines all of the $f_\alpha(t)$ uniquely. If c is the constant term of f, we define $f_0(t) \equiv c$.

Thus we have projected f on a complete set of mutually orthogonal invariant closed linear manifolds and have made each projection correspond to a function in $L_2(0,1)$. The functions $\{f_\alpha(t)\}$, (all α), and the function $f_0(t)$, will be said to *determine* f.

THEOREM 2.1. *If f is determined by the functions $\{f_\alpha(t)\}$ and $f_0(t)$, $E(\lambda)f$ for $0 < \lambda < 1$ is determined by $\phi_\lambda(t) \cdot f_\alpha(t)$ and $f_0(t)$, where $\phi_\lambda(t)$ is given by* (2.6).

[10] $L_2(0,1)$ is the class of all complex valued functions of the real variable t, $0 \leq t \leq 1$, which are measurable and integrable squared.

Proof. The fact that $f_0(t)$ is the same for f and $E(\lambda)f$ follows at once from the fact that U leaves the constant term in the Fourier series of f unchanged.

The function f less its projection on $\mathfrak{M}_0$ can be written as a sum of components where each component is the projection of f on some one of the mutually orthogonal invariant $\mathfrak{M}_\alpha$. It is sufficient to prove the theorem for each component separately. We shall give the proof for functions in $\mathfrak{M}_1$ where $\mathfrak{M}_1 = \mathfrak{M}_\alpha$ with $\alpha = (1)$. The proof for the other manifolds is the same.

Let f be a function in $\mathfrak{M}_1$. Then it has the Fourier series

$$\sum_j b_j e^{2\pi i x_j(\omega)} \tag{2.11}$$

where

$$b_j = \int_0^1 e^{-2\pi i j t} f_1(t)\, dt. \tag{2.12}$$

Applying (2.4) [11] we see that

$$E(\lambda) = \text{l. i. m.}_{\mu\to\infty} \sum_{m=-\mu}^{m=+\mu} \Big[\sum_j a_m(\lambda) b_j e^{2\pi i x_{j+m}(\omega)}\Big], \tag{2.13}$$

for $0 < \lambda < 1$. Thus $E(\lambda)f$ is also in $\mathfrak{M}_1$:

$$E(\lambda)f = \sum_j c_j e^{2\pi i x_j(\omega)}. \tag{2.14}$$

Then by Parseval's identity applied to $\phi_\lambda(t)$ and $f_1(t)e^{-2\pi i j t}$,

$$c_j = \int_0^1 \phi_\lambda(t) f_1(t) e^{-2\pi i j t} dt, \tag{2.15}$$

which is the desired result.

THEOREM 2.2. *For f and g in $L_2(\Omega^*)$, $(E(\lambda)f, g)$ is absolutely continuous after a jump at $\lambda = 0$ has been removed.*

Proof. We shall first prove that for $f \in L_2(\Omega^*)$ and $(f, 1) = 0$, $(E(\lambda)f, f)$ is absolutely continuous. Then, since, for $c =$ constant,

$$(E(\lambda)[f + c], [f + c]) = \begin{cases} (E(\lambda)f, f) + |c|^2 \text{ for } \lambda > 0, \\ 0 \text{ for } \lambda \leq 0, \end{cases} \tag{2.16}$$

it will follow that $(E(\lambda)[f + c], [f + c])$ minus the function of λ which is zero for $\lambda \leq 0$ and $|c|^2$ elsewhere is absolutely continuous. The result for

[11] It is well known that the transformation U has no characteristic functions other than the constants, that is that the shift is metrically transitive and has no angle variables. (For a proof following the methods of this paper cf. J. L. Doob, *Transactions of the A. M. S.*, vol. 36 [1934], pp. 761-763.) Hence $E(\lambda)$ is continuous for $\lambda > 0$.

$(E(\lambda)f, g)$, where f and g are arbitrary elements in $L_2(\Omega^*)$, will follow from the above by the usual considerations of $f+g$ and $f+ig$.

Let $\{f_a(t)\}$ be the functions which determine f. Then by Parseval's identity

$$(2.17) \qquad (E(\lambda)f, f) = \sum_a \int_0^\lambda |f_a(t)|^2\, dt, \qquad 0 < \lambda < 1.$$

Since $(E(\lambda)f, f)$ is finite, the series $\sum_a |f_a(t)|^2$ converges for almost all t to an integrable function $\bar{f}(t)$ and

$$(2.18) \qquad (E(\lambda)f, f) = \sum_a \int_0^\lambda |f_a(t)|^2\, dt = \int_0^\lambda \bar{f}(t)\, dt,$$

for $0 < \lambda < 1$. But both sides of (2.18) are continuous at $\lambda = 0$ and $\lambda = 1$. Hence (2.18) holds for $0 \leq \lambda \leq 1$ and this completes the proof.

3. Unitary operators arising from general measure preserving transformations. In this section we consider an abstract space Ω with a measure P of Lebesgue type—that is, completely additive and non-negative on some Borel field of Ω—with $P(\Omega) = 1$; that is, we consider a probability space. Let U be a unitary operator in $L_2(\Omega)$. Motivated by the results of Section **2** we study manifolds determined by $\{U^n\phi\}$, $n = 0, \pm 1, \pm 2, \cdots$, for some given $\phi \in L_2(\Omega)$. An important consideration in the preceding section was that the sequence (2.9) was an independent sequence. The functions $a(\omega)$, $b(\omega)$, $\cdots, h(\omega)$ are said to be independent if

$$(3.1) \quad P\{a \in A, b \in B, \cdots, h \in H\} = P\{a \in A\}, P\{b \in B\} \cdots P\{h \in H\}\ ^{12}$$

for every collection $A, B, \cdots, H$ of Borel sets in the complex plane. An infinite sequence of functions is said to be an *independent sequence* if every finite subset is independent. In what follows we shall deal with independent functions and shall make use of the well known theorem that if f and $g \in L_2(\Omega)$ are independent, $(f, g) = (f, 1) \cdot (1, g)$.

THEOREM 3.1. *If U is a unitary operator in $L_2(\Omega)$ and if $\mathfrak{M}$ is the closed linear manifold determined by a sequence of functions $\{\phi_j\}$, $j = 0, \pm 1, \pm 2, \cdots$ such that $U\phi_j = \phi_{j+1}$ and $(\phi_j, \phi_k) = 0$ for $j \neq k$, then, for $f, g \in \mathfrak{M}$, $(E(\lambda)f, g)$ is absoutely continuous, and*

$$(3.2) \qquad (E(\lambda)\phi_j, \phi_j) = \lambda(\phi_j, \phi_j), \qquad 0 \leq \lambda \leq 1.$$

---

[12] $\{c\}$ will stand for the ω-set where the condition c holds. These brackets will also indicate sequences but the context will prevent confusion.

Proof. Suppose that $f \in \mathfrak{M}$. To apply (2.4) we must show that $(E(\lambda)f, f)$ is continuous. In order to do this we shall show that f is orthogonal to the manifold of invariant and characteristic functions: If $U_g = e^{2\pi i\lambda}g$, then $(g, \phi_n) = (Ug, U\phi_n) = e^{2\pi i\lambda}(g, \phi_{n+1})$. Hence the Fourier coefficients (g, ϕ_n) have the same modulus and so by Bessel's inequality they all vanish: $(f, g) = 0$.

Since f is in $\mathfrak{M}$ it can be written

$$f = \sum_j c_j\phi_j, \text{ where } \sum_j |c_j|^2 < \infty. \tag{3.3}$$

Then by (2.4)

$$\begin{aligned}(E(\lambda)f, f) &= \Big(\sum_m \sum_j a_m(\lambda)c_j\phi_{j+m}, \sum_j c_j\phi_j\Big) \\ &= (\phi_0, \phi_0) \sum_m \sum_j a_m(\lambda)c_j\bar{c}_{j+m},\end{aligned} \tag{3.4}$$

for $0 < \lambda < 1$. But, since $\sum_j |c_j|^2$ is finite, the c_j are the Fourier coefficients of a function $F(t) \in L_2(0, 1)$. Hence the c_{j+m}, for fixed m, are the Fourier coefficient of $e^{-2\pi imt}F(t)$. Then, by Parseval's identity,

$$(E(\lambda)f, f) = (\phi_0, \phi_0) \int_0^\lambda |F(t)|^2\, dt, \tag{3.5}$$

for $0 < \lambda < 1$. But both sides of (3.5) are continuous at $\lambda = 0$ and at $\lambda = 1$. Hence the inequality holds for $0 \leq \lambda \leq 1$.

This proves the first part of the theorem for $f = g$. The extension to $f \neq g$ follows as in the proof of Theorem 2.3.

To prove (3.2) set $f = \phi_j$ in the above argument. Then $F(t) = e^{-2\pi ijt}$ and (3.2) follows since $(\phi_0, \phi_0) = (\phi_j, \phi_j)$.

THEOREM 3.2. *If U is a unitary operator in $L_2(\Omega)$, a necessary and sufficient condition that*

$$(U^n f, g) = (f, 1)(1, g), \tag{3.6}$$

for $n \neq 0$, is that

$$(E(\lambda)f, g) = \lambda(f, g) + (1 - \lambda)(f, 1)(1, g), \tag{3.7}$$

for $0 < \lambda \leq 1$.

Since $U(1) = 1$, $(f, 1)(1, g) = (U^n f, 1)(1, g)$, and hence (3.6) implies that $U^n f$ and g are uncorrelated for $n \neq 0$.

Proof. Suppose that (3.7) holds. Then

$$(3.8)\qquad (U^n f, g) = \int_0^1 e^{2\pi i n\lambda} d(E(\lambda)f, g)$$
$$= [(f, g) - (f, 1)(1, g)] \int_0^1 e^{2\pi i n\lambda} d\lambda + (f, 1)(1, g)$$
$$= (f, 1)(1, g),$$

for $n \neq 0$, which proves sufficiency.

Suppose that (3.6) holds. Then

$$(3.9)\qquad (U^n f, g) = \int_0^1 e^{2\pi i n\lambda} d(E(\lambda)f, g) = \begin{cases} (f, g) \text{ for } n = 0, \\ (f, 1)(1, g) \text{ for } n \neq 0. \end{cases}$$

This determines $(E(\lambda)f, g)$ except for an additive constant. But $(E(\lambda)f, g)$ as given by (3.7) satisfies (3.9). Hence (3.7) is correct except possibly for an additive constant. However, if we let λ increase to one in (3.7), we see that the constant is zero, which proves necessity.

It is interesting to note that if the statement "$n \neq 0$" had been omitted from the hypothesis of this theorem, the result would have been

$$(3.10)\qquad (E(\lambda)f, g) = (f, 1)(1, g), \qquad 0 < \lambda \leqq 1.$$

THEOREM 3.3. *Let U be a unitary operator in $L_2(\Omega)$ arising from a measure preserving transformation T on Ω and consider the sequence $\{\phi_n(\omega)\}$, $n = 0, \pm 1, \pm 2, \cdots$, where $U\phi_n = \phi_{n+1}$. Denote by $\mathfrak{N}$ the closed linear manifold determined by all bounded Baire functions of a finite number of the ϕ_n that is functions $f(\phi_1, \cdots, \phi_\nu)$, where $f(x_1, \cdots, x_\nu)$ is a bounded Baire function of the complex arguments $x_1, \cdots, x_\nu$. Then, if the ϕ_n form an independent set, $(E(\lambda)f, g)$ for $f, g \in \mathfrak{N}$ is absolutely continuous after a jump at $\lambda = 0$ has been removed. Furthermore*

$$(3.11)\qquad (E(\lambda)\phi_j, \phi_j) = \lambda(\phi_j, \phi_j) + (1 - \lambda)\,|\,(\phi_j, 1)\,|^2, \qquad 0 < \lambda \leq 1.$$

Proof. The proof of the first part of the theorem will be similar to the proof of Theorem 2.2. We shall construct a complete o. n. set in $\mathfrak{N}$ having the properties of (2.9); that is, we shall decompose $\mathfrak{N}$ into a complete set of mutually orthogonal invariant linear manifolds on each of which there is a complete o. n. set $\{\psi_n(\omega)\}$, $n = 0, \pm 1, \pm 2, \cdots$, such that $U\psi_n = \psi_{n+1}$. The rest of the proof of this part of the theorem will then follow the proof of Theorem 2.2 so closely that we shall omit it.

Consider the closed linear manifold in $\mathfrak{N}$ determined by functions of ϕ_0 only. By the usual approximation and orthogonalization processes we can find a complete o. n. set in this manifold. We can suppose that 1 appears in this

o. n. set and can denote the set by: $1, \psi_n^{(0)}(\omega)$. The $\psi_n^{(0)}$ are, of course, bounded Baire functions of ϕ_0. If in the above argument we replace ϕ_0 by ϕ_j, $j = \pm 1, \pm 2, \cdots$, we get a complete o. n. set: $1, \psi_n^{(j)}(\omega)$, $n = 0, \pm 1, \pm 2, \cdots$, in the manifold of functions of ϕ_j. $\psi_n^{(j)}$ is the same function of ϕ_j that $\psi_n^{(0)}$ is of ϕ_0. Hence, since $U^j\phi_0 = \phi_j$, $U^j\psi_n^{(0)} = \psi_n^{(j)}$.

The complete o. n. set similar to (2. 9) that we seek will consist of 1 and all finite products of $\psi_n^{(j)}$'s such that no $\psi_n^{(j)}$ appears more than once in any product. To prove that this set has the desired properties it is sufficient to establish the following three results:

(a) If two of our products have the same factors, their inner product is one; otherwise they are orthogonal.

(b) The set: $1, \{\psi_{n_1}^{(j_1)} \cdots \psi_{n_k}^{(j_k)}\}$, $j_1, \cdots, j_k = 0, \pm 1, \pm 2, \cdots$, $k = 1, 2, \cdots, p$ is dense in the closed linear manifold of bounded Baire functions of $\phi_{n_1}, \cdots, \phi_{n_p}$.

(c) $U[\psi_{n_1}^{(j_1)} \cdots \psi_{n_k}^{(j_k)}] = \psi_{n_1}^{(j_1+1)} \cdots \psi_{n_k}^{(j_k+1)}$.

To prove (a) consider the integral of $\psi_{m_1}^{(j_1)} \cdots \psi_{m_p}^{(j_k)} \cdot \overline{\psi_{n_1}^{(k_1)} \cdots \psi_{n_q}^{(k_q)}}$. It is of the form $\int_\Omega |\psi_{s_1}^{(l_1)}|^{\delta_1} \cdots |\psi_{s_r}^{(l_r)}|^{\delta_r} dP = \int_\Omega |\psi_{s_1}^{(l_1)}|^{\delta_1} d_1 P \cdots \int_\Omega |\psi_{s_r}^{(l_r)}|^{\delta_r} dP$, where the δ's are one or two and if one of them is one, there is no absolute value sign on the corresponding term. Each integral on the right is one if its δ is two and zero if its δ is one. But all the δ's are two if and only if the two products we are considering have the same factors.

Statements (b) and (c) follow from the way in which the $\psi_n^{(j)}$ were defined and the fact that U arose from a measure preserving transformation. The first part of the proof of the theorem now follows the proof of Theorem 2. 2.

To prove the second part of the theorem it is only necessary to observe that the sequence $\{\phi_j\}$ satisfies the hypotheses of Theorem 3. 2. Then (3. 11) is a result of applying (3. 7) to this sequence.

Suppose now that U is a unitary operator in $L_2(\Omega)$ arising from a measure preserving transformation T in Ω and that $\{U^n f\}$, $n = 0, \pm 1, \pm 2, \cdots$, is a sequence of functions independent in pairs. Then if $\phi(y)$ and $\psi(y)$ are two Baire functions such that $\phi(f)$ and $\psi(f)$ are in $L_2(\Omega)$, $\psi(f)$, and $\phi(U^n f)$ are independent for $n \neq 0$. But $\phi(U^n f) = \phi(f(T^n\omega)) = U^n\phi(f)$. Then $U^n\phi(f)$ and $\psi(f)$ are independent and hence uncorrelated for $n \neq 0$. Consequently by Theorem 3. 2

$$(3.12) \quad (E(\lambda)\phi(f), \psi(f)) = (1-\lambda)(\phi(f), 1)(1, \psi(f)) + \lambda(\phi(f), \psi(f)), \lambda \neq 0.$$

Let $\psi(y)$ be the characteristic function of the Borel set A in the complex plane. Then (3.12) becomes

$$(3.13)\qquad \int_{\{f \varepsilon A\}} E(\lambda)\phi(f)dP = (1-\lambda)\mathcal{E}[\phi(f)]\cdot P\{f \varepsilon A\} + \lambda \int_{\{f \varepsilon A\}} \phi(f)dP, \lambda \neq 1$$

By the definition of conditional expectation [14] it follows that

$$(3.14)\qquad \mathcal{E}[f = \xi, E(\lambda)\phi(f)] = (1-\lambda)\mathcal{E}[\phi(f)] + \lambda\phi(\xi), \lambda \neq 0.$$

But the above discussion is also reversible. Hence we have

THEOREM 3.4. *If U arises from a measure preserving transformation in Ω a necessary and sufficient condition that $\{U^n f\}$, $n = 0, \pm 1, \pm 2, \cdots$, be a sequence of functions independent in pairs is that*

$$(3.15)\qquad \mathcal{E}[f = \xi, E(\lambda)\phi(f)] = (1-\lambda)\mathcal{E}[\phi(f)] + \lambda\phi(\xi),$$

for $\lambda = 0$, and for every Baire function ϕ such that $\phi(f)$ belongs to $L_2(\Omega)$.

UNIVERSITY OF ILLINOIS,
PURDUE UNIVERSITY.

[13] $\mathcal{E}[h(\omega)] = \int_\Omega h(\omega)dP$ is called the *expectation* of $h(\omega)$.

[14] $\mathcal{E}[g = \xi, h(\omega)]$ is the *conditional expectation* of $h(\omega)$ for $g(\omega) = \xi$. For the definition and discussion of this concept cf. Kolmogoroff, *loc. cit.*, pp. 41-50.

INTEGRABILITY IN THE LARGE AND DYNAMICAL STABILITY.*

By PHILIP HARTMAN and AUREL WINTNER.

1. The most fundamental but also the most intricate of the various definitions of a stable solution of a conservative dynamical system may roughly be expressed by the requirement that small changes in the initial conditions should produce small changes in the whole solution. The analytical difficulties involved by this classical condition are due to the fact that continuous dependence on the initial conditions is required uniformly for the whole time axis.

If the solution considered is a solution of equilibrium, it is or is not stable in this sense, according as it does or does not satisfy the Poincaré-Birkhoff criterion; a criterion requiring the existence of invariant sets closing down on the point of equilibrium.

If the solution is not a solution of equilibrium, the only stable cases known today are of the trivial type of Liouville, as exemplified by the Diophantine cases of the three integrable tops or of the problem of geodesics on surfaces of revolution. These dynamical systems are integrable not only in the sense that the full amount of (non-local, analytic) integral hypersurfaces exists, but also in the sense that the Diophantine analysis of the solution curves on these hypersurfaces supplies for the solutions actual Fourier series, and not merely formal expansions.

The results of the present paper will imply that these properties of a system of the Liouville type are interdependent; so much so that their simultaneous appearance has nothing to do with the elementary nature of the dynamical system at hand. In fact, it turns out that integrability in the sense of anharmonic Fourier series of the type in question is equivalent to stability in the classical sense referred to above; it being understood that stability is meant with reference to solutions on the relevant hypersurfaces or invariant sets of the phase space.

2. The theorem to be proved supplies a criterion of stability in case the solution considered is regionally transitive; a case in which the situation is precisely opposite to the one covered by the Poincaré-Birkhoff criterion, where the solution path degenerates to a point. Correspondingly, the criterion to be proved explains, to some extent, the apparent scarcity of dynamical systems

* Received September 29, 1941.

possessing solutions which are stable but do not represent equilibria. Needless to say, the result is irrelevant in the case of strictly periodic solutions; the latter being always reducible to equilibrium points, as shown by the classical device of a local surface transformation.

It is of methodical importance that the flow representing the general solution of the conservative system will not be postulated to be of the dynamical type, not even in the sense of mere incompressibility. This situation is relevant from the point of view of a by-product of the main result to be proved. For it will thus be possible to obtain as a corollary a result similar to Poincaré's theorem on those (orientation-preserving) homeomorphisms of a circle for which no iterate has a fixed point. In fact, starting with the purely topological assumptions of regional transitivity and stability, the corollary supplies the existence of an intrinsic invariant measure with reference to which metrical transitivity takes place.

3. Let a compact set S of points $x = (x_1, \cdots, x_n)$ be an invariant set of $dx/dt = X(x)$, where $X = (X_1, \cdots, X_n)$ is, for instance, of class C^1 on S. The dimension number of S need not exist, and if it exists, it need not be n. The incompressibility condition, $\operatorname{div}_x X(x) \equiv 0$, of the flow defined by $dx/dt = X(x)$ on S is not assumed. Let $x(t) = x(t, x_0)$ denote the solution which reduces for $t = 0$ to the point x_0 of S. Since S is a compact invariant set, the path $x(t)$ exists for $-\infty < t < \infty$ and its closure is contained in S.

If the closure of the path is identical with S, that is, if the curve $x = x(t), -\infty < t < \infty$, is dense on S, the solution $x(t)$ is called regionally transitive on S. (For sake of brevity, the "alpha" and "omega" cases of regional transitivity will not be considered).

A solution will be called almost periodic if it is almost periodic in the sense of Bohr.

The stability of a given solution on S will be meant, with reference to all solutions on S, in the sense corresponding to the classical Minding-Dirichlet definition of a stable equilibrium. Accordingly, a solution $x(t) = x(t, x_0)$ will be called stable on S if there exists for every $\epsilon > 0$ a $\delta = \delta_\epsilon(x_0)$ such that the deviation, $|y(t) - x(t)|$, of an arbitrary solution, $y(t)$, from the given solution, $x(t)$, is less than ϵ for $-\infty < t < \infty$ whenever it is less than $\delta_\epsilon(x_0)$ for a single value of t.

It will be shown in Sections **6** and **7** that *a solution regionally transitive on S is stable on S if and only if it is almost periodic; in which case every solution on S is almost periodic.*

This theorem implies that if one solution is regionally transitive and stable on S, the same is true of all solutions on S. The proof of the theorem will

depend on two uniformity statements which will be proved in Sections **4** and **5** and can be formulated as follows:

The stability of the solutions must be uniform on S, in the sense that $\delta_\epsilon(x_0)$ can be chosen independent of $x_0 = x(0)$ for every fixed $\epsilon > 0$. This uniformity in stability is paralleled by a corresponding uniformity in almost periodicity, in the sense that *the translation numbers belonging to any fixed $\epsilon > 0$ are independent of the integration constant $x_0 = x(0)$ of the path on S.*

4. It will first be shown that, if one solution is stable and regionally transitive on S, the general solution, $x(t, x_0)$, of $dx/dt = X(x)$ is a uniformly continuous function of the position on the product space of T and S_0, where T denotes the infinite t-axis and S_0 the space of all initial positions $x_0 = x(0)$ on S. To this end, it is sufficient to show that all solutions on S are stable on S and that their stability is uniform in the sense explained at the end of Section **3**.

Let $\delta_\epsilon(x_0)$ denote the δ belonging to the given stable solution $x(t)$ on S and to an $\epsilon > 0$. Since $x(t)$ is supposed to be regionally transitive on S, there exists for every point, say y_0, of S a date, say $t' = t'(\epsilon, x_0; y_0)$, satisfying $|x(t') - y_0| < \frac{1}{2}\delta_\epsilon(x_0)$. Let z_0 be any point of S such that $|z_0 - y_0| < \frac{1}{2}\delta_\epsilon(x_0)$. In order to prove the uniform stability of all solutions on S, it will be sufficient to show that the solutions $y(t)$, $z(t)$ determined by the respective initial conditions $y(0) = y_0$, $z(0) = z_0$ satisfy the inequality $|y(t) - z(t)| < 2\epsilon$ for $-\infty < t < \infty$.

It is clear from the choice of t' and z_0 that $|x(t') - y_0| < \delta_\epsilon(x_0)$ and $|x(t') - z_0| < \delta_\epsilon(x_0)$. Hence, $|x(t + t') - y(t)| < \delta_\epsilon(x_0)$ and $|x(t + t') - z(t)| < \delta_\epsilon(x_0)$ are satisfied by a certain t; namely, by $t = 0$. Since $\delta_\epsilon(x_0)$ belongs to the stable solution $x(t)$, and since the latter represents for $-\infty < t < \infty$ the same path as does the solution $x(t + t')$, where t' is fixed, it follows from the definition of $\delta_\epsilon(x_0)$, that $|x(t + t') - y(t)| < \epsilon$ and $|x(t + t') - z(t)| < \epsilon$ for every t. But this implies that, as stated above, $|y(t) - z(t)| < 2\epsilon$ holds for every t.

5. In order to prove the second of the uniformity statements formulated at the end of Section **3**, suppose that there exists on S one regionally transitive solution which is almost periodic. The assertion is that all solutions on S are almost periodic and have common translation numbers for every fixed $\epsilon > 0$.

Let $x(t)$ be the given almost periodic solution. Since it is supposed to be regionally transitive on S, there exists to every point, say y_0, of S a sequence of dates, say $\{t_k\}$, such that $x(t_k) \to y_0$. According to the compactness criterion for the almost periodicity of $x(t)$, the sequence $\{t_k\}$ contains a subsequence,

say $\{t'_k\}$, such that the corresponding translated sequence, $\{x(t+t'_k)\}$, of $x(t)$ is uniformly convergent for $-\infty < t < \infty$. Let $y(t)$ denote the limit function of this translated sequence.

Since $x(t_k) \to y_0$ implies that $x(t'_k) \to y_0$, it is clear from the local uniqueness and continuity theorem of ordinary differential equations, that the function $y(t)$, $-\infty < t < \infty$, is identical with the solution satisfying the initial condition $y(0) = y_0$. But y_0 was chosen as an arbitrary point of S. Hence, all that remains to be shown is that every translation number of $x(t)$ belonging to an ϵ is a translation number of $y(t)$ belonging to the same ϵ. But this is clear from the fact that $x(t+t'_k)$ tends to $y(t)$ uniformly for $-\infty < t < \infty$, as $k \to \infty$.

In view of Bohr's theory of translation classes, the above result implies that, if $\sum a_m e^{i\lambda_m t}$, where $a_m \neq 0$, is the Fourier series of $x(t)$, the amplitudes, $|a_m|$, and the frequencies, λ_m, are independent of $x_0 = x(0)$ on S; so that only the phases, $\arg a_m$, vary with the integration constants determining a solution $x(t)$ on S.

6. In order to prove the necessary and sufficient condition announced in Section **3**, suppose first that a regionally transitive solution, $x(t)$, is stable on S.

According to Section **4**, all solutions on S must then be uniformly stable. This means that there exists for every $\epsilon > 0$ an $\eta_\epsilon > 0$ such that the deviation, $|z(t) - y(t)|$, of any two solutions on S is less than ϵ for every t if it is less than η_ϵ for a single t. Consequently, if $\{t'_k\}$ is any sequence of dates such that the sequence, $\{x(t'_k)\}$, of the corresponding points of the path $x(t)$ tends to a point, say y_0, then the sequence $\{x(t+t'_k)\}$ tends to $y(t)$ uniformly for $-\infty < t < \infty$, where $y(t)$ denotes the solution determined by the initial condition $y(0) = y_0$.

Now let $\{t_k\}$ be any sequence of dates. Since S is compact and contains the path $x(t)$, one can select from $\{t_k\}$ a subsequence, say $\{t'_k\}$, such that the corresponding points of the path form a convergent sequence, $\{x(t'_k)\}$. But this was seen to imply that the sequence $\{x(t+t'_k)\}$ is uniformly convergent for $-\infty < t < \infty$.

Accordingly, $x(t)$ has the property that every sequence, $\{x(t+t_k)\}$, of its translated functions contains a uniformly convergent subsequence, $\{x(t+t'_k)\}$. It follows therefore from the compactness criterion of almost periodicity that $x(t)$ is almost periodic.

7. In order to prove the converse, suppose that a regionally transitive solution, $x(t)$, on S is almost periodic. This implies, by Section **5**, that there

exists for every $\epsilon > 0$ an $l_\epsilon > 0$ such that every t-interval of length l_ϵ contains a translation number τ_ϵ satisfying $|y(t+\tau_\epsilon) - y(t)| < \epsilon$ for every t and for every solution $y(t)$ on S.

Since S is compact, it is clear from the local continuity theorem of ordinary differential equations, that there exists for every $\epsilon > 0$ and for every $L > 0$ a positive $\eta = \eta(\epsilon, L)$ such that the deviation, $|x(t) - y(t)|$, of two arbitrary solutions on S is less than ϵ on the t-interval $t_0 - L < t < t_0 + L$, if it is less than $\eta(\epsilon, L)$ for $t = t_0$, where t_0 is arbitrary.

Choose $L = l_\epsilon$ and put $\eta_\epsilon = \eta(\epsilon, l_\epsilon)$. Thus $|x(t) - y(t)| < \epsilon$ holds on the t-interval $t_0 - l_\epsilon < t < t_0 + l_\epsilon$ (of length $2l_\epsilon$) whenever $|x(t_0) - y(t_0)| < \eta_\epsilon$. But every t-interval of length l_ϵ contains a translation number τ_ϵ such that $|y(t+\tau_\epsilon) - y(t)| < \epsilon$ holds for $-\infty < t < \infty$ and for every solution $y(t)$ on S. In particular $|x(t+\tau_\epsilon) - x(t)| < \epsilon$ for $-\infty < t < \infty$.

The last three ϵ-estimates obviously imply that, if $|x(t) - y(t)|$ is less than η_ϵ for a single t, then it is less than 3ϵ for $-\infty < t < \infty$. But this means that $x(t)$ is stable on S.

8. This completes the proof of all the statements made in Section **3.**

In the sequel, use will be made of the fact that every almost periodic solution $x(t)$ on S possesses a distribution function $\phi = \phi(E)$, in the following sense: ϕ is a non-decreasing, additive set-function, defined for every Borel subset, E, of S in such a way that, unless E is a discontinuity set of the monotone set-function ϕ, the ratio of $(u, v)_E$ to the length, $v - u$, of the t-interval $u < t < v$ tends to the limit $\phi(E)$ as $v - u \to \infty$, where $(u, v)_E$ denotes the linear measure of the set of points t of the t-interval $u < t < v$ along which the point $x(t)$ of S is in E.

Clearly, the distribution function of every almost periodic solution $x(t)$ on S defines on S a Lebesgue measure such that $\phi(S) = 1$. However, this Lebesgue measure depends, in general, on the initial condition $x(0) = x_0$ determining the solution $x(t)$; in which case it is not an invariant measure of the flow defined by $dx/dt = X(x)$.

But suppose that there exists a stable solution, $x(t)$, which is regionally transitive on S. Then, according to Section **3**, every solution on S (in particular $x(t)$ itself) is almost periodic; moreover, by the proof in Section **5**, there exists to every solution, $y(t)$, a sequence, $\{t_k\}$, such that the function $x(t + t_k)$ of t tends to $y(t)$ uniformly for $-\infty < t < \infty$, as $k \to \infty$. This obviously implies that $y(t)$, and therefore every solution on S, has the same distribution, say $\phi = \phi(E)$, as $x(t)$.

It follows that ϕ, when thought of as a Lebesgue measure on S, is invariant under the flow defined by the general solution of $dx/dt = X(x)$ on S. Since

$\phi(E)$ is the common distribution function of all these solutions, it is clear that *the flow is metrically transitive on* S with reference to the invariant measure ϕ on S. Needless to say, metrical transitivity and ergodicity are equivalent notions in the present case.

9. Birkhoff's ergodic theorem is equivalent to the assertion that almost all paths have a distribution function whenever the flow on S has an invariant measure for which S is of finite measure. It is implied by the geodesic problems depending on Fuchsian groups, that the zero set excluded by Birkhoff's theorem can be non-enumerable even in case the flow is metrically transitive.

Thus it is of interest that the zero set is vacuous in the case of Section **8**. In this respect, the situation in the general case covered by Section **8** is therefore the same as in the case of a Kronecker-Weyl flow on a torus; a case on which the elementary examples of the Liouville type depend (cf. Section **1**).

It would even be possible that this parallelism is not a mere coincidence. In order to analyze the situation, it would be necessary to find the topological invariants of flows containing paths which are regionally transitive and stable. But such a task could hardly be attacked today, even if S is restricted to be a manifold.

Thus it is of particular interest that the facts formulated in Section **3** hold without the explicit integrability assumptions implied by the elementary problems mentioned at the beginning of Section **1**.

ARMY OF THE UNITED STATES,
THE JOHNS HOPKINS UNIVERSITY.

REFERENCES

Birkhoff, G. D., *Dynamical Systems*, New York, 1927.

Franklin, Ph., "Almost periodic recurrent motions," *Mathematische Zeitschrift*, vol. 39 (1929), pp. 325-331.

Haviland, E. K., "On statistical methods in the theory of almost periodic functions," *Proceedings of the National Academy of Sciences*, vol. 19 (1933), pp. 549-555.

Wiener, N., and Wintner, A., "On the ergodic dynamics of almost periodic systems," *American Journal of Mathematics*, vol. 58 (1941), pp. 794-824.

Wintner, Aurel, *The Analytical Foundations of Celestial Mechanics*, Princeton, 1941.

THE DISCRETE CHAOS.*

By Norbert Wiener and Aurel Wintner.

Introduction

A characteristic difficulty of problems requiring the introduction of a measure with infinitely many dimensions is that ratios which are normally finite become either infinite or zero in the majority of cases. Such difficulties are of a trivial nature in case of complete independence of infinitely many random variables, since in this case the proper measure is simply the product measure. However, most of the physical theories presuppose measure which belongs to infinitely many dimensions but not to the trivial case of a product measure.

In classical mechanics, the Maxwellian picture is that of a large number of individual particles, and averages are taken when this number becomes infinite; while in the picture of Boltzmann and Gibbs, a system with a fixed finite degree of freedom is studied as a statistical function of its initial phase determination. The standard case of a Maxwellian description is that in which the particles have a vanishingly small coupling; so that the measure which determines probabilities reduces to a product measure in the limiting case. On the other hand, the standard formulation of the Boltzmann-Gibbs statistics involves densities which cease to exist when the degree of freedom is *actually* infinite. Nevertheless, notions like " number of distinct phases " or statements like that of the " almost certain increase of entropy " are meaningless or untrue with reference to the probability measure of a system with a *fixed* finite degree of freedom.

Thus there arises the question as to the existence of a definite probability measure in terms of which it is possible to unite the advantages, and eliminate the mathematical inadequacies, of the models of Maxwell and Boltzmann-Gibbs.

Although the present paper does not deal with statistical mechanics, the statistical problems to be considered are characterized by difficulties of the type just described. Of course, the problem will not be the postulation of a suitable measure, without which statistical statements are meaningless, but rather an existence proof, which must necessarily be based on an explicit construction in terms of a finite number of random variables.

* Received September 24, 1941.

PART I

The Distribution Functions

1. Let $\mathcal{E}$ be a field of sets E which is closed with respect to finite or enumerable logical addition and multiplication. Let $\mathcal{E}^{(n)}$ denote the field which is the closure (with respect to addition and multiplication) of the product $\mathcal{E}^n = \mathcal{E}^{n-1} \times \mathcal{E}$, where $\mathcal{E}^1 = \mathcal{E}$. It will be convenient to let $\mathcal{E}^{(0)}$ (and $\mathcal{E}^0$) denote the field consisting of the empty set. The logical sum of all sets E will be denoted by S, provided S is an E.

Let there be given, for every n, a non-negative, additive function Ψ_n of the sets which constitute $\mathcal{E}^{(n)}$. For those sets in $\mathcal{E}^{(n)}$ which are in $\mathcal{E}^n$, i. e., which are of the form $E_1 \times \cdots \times E_n$, the function Ψ_n can be written as a function $\Psi_n = \Psi_n(E_1, \cdots, E_n)$ of n sets E_j each of which is in $\mathcal{E}$. It will be assumed that $\Psi_n(E_1, \cdots, E_n)$ is a symmetric function of the n variables E_j if $n \geqq 1$, while $\Psi_0 = 1$.

Let $E_1, \cdots, E_k$ be mutually disjoint sets of $\mathcal{E}$. For every integer $k \geqq 0$ and for every k-uple, $(n_1, \cdots, n_k)$, of integers $n_j \geqq 0$, let

$$\Phi_{n_1 \ldots n_k}(E_1, \cdots, E_k) = \Psi_n(E_1, \cdots, E_1, \cdots, E_k, \cdots, E_k), \quad \text{where } n = n_1 + \cdots + n_k; \tag{1}$$

it being understood that every E_j occurs n_j times in Ψ_n and that sets with different indices are disjoint. In particular

$$\Psi_k(E_1, \cdots, E_k) = \Phi_{1 \ldots 1}(E_1, \cdots, E_k), \text{ where } 1 + \cdots + 1 = k. \tag{1 bis}$$

Hence, the assignment of the functions Ψ_n is equivalent to the assignment of the functions $\Phi_{n_1 \ldots n_k}$. Since $\Psi_0 = 1$, the Φ which has no subscript ($k = 0$) is 1. It is important that (1) is undefined unless $E_1, \cdots, E_k$ are mutually disjoint.

It will be convenient to assume that, for every k and arbitrary E_j,

$$\Phi_{n_1 \ldots n_k}(E_1, \cdots, E_k) < c^{n_1 + \ldots + n_k}, \quad \text{where} \quad c = c(E_1, \cdots, E_k), \tag{2}$$

holds for a sufficiently large c which is independent of the n_j. This is the case if and only if

$$f_{E_1 \ldots E_k}(z_1, \cdots, z_k) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_k=0}^{\infty} \frac{(z_1 - 1)^{n_1} \cdots (z_k - 1)^{n_k}}{n_1! \cdots n_k!} \Phi_{n_1 \ldots n_k}(E_1, \cdots, E_k) \tag{3}$$

is an entire function of exponential type in the k complex variables z_j, where k and $E_1, \cdots, E_k$ are arbitrary. It is clear from (1) and from the symmetry of $\Psi_n(E_1 \cdots, E_n)$ in the E_j, that (3) is invariant under any simultaneous permutation of the z_j and the E_j. The f which has no subscript ($k = 0$) is the constant 1. Furthermore,

$$f_{E_1 \ldots E_k}(1, \cdots, 1) = 1. \tag{4}$$

More generally, (3) and (1) imply that

$$(5) \qquad f_{E_1 \ldots E_{k-1}}(z_1, \cdots, z_{k-1}) = f_{E_1 \ldots E_k}(z_1, \cdots, z_{k-1}, 1),$$

since the Φ which has no index is 1. It is similarly verified that, in virtue of the binomial theorem and the additivity of the set functions Ψ, the relation

$$(6) \qquad f_{E_0 E_1 E_2 \ldots E_k}(z_1, z_1, z_2, , \cdots, z_k) = f_{E_0 + E_1 E_2 \ldots E_k}(z_1, z_2, \cdots, z_k)$$

holds for any pair of disjoint sets E_0, E_1 of $\mathcal{E}$.

2. Reorder (3) according to the powers of the z_j, obtaining

$$(7) \qquad f_{E_1 \ldots E_k}(z_1, \cdots, z_k) = \sum_{m_1=0}^{\infty} \cdots \sum_{m_k=0}^{\infty} \phi_{m_1 \ldots m_k}(E_1, \cdots, E_k) z_1^{m_1} \cdots z_k^{m_k}$$

as the definition of the ϕ's. It is clear from the corresponding properties of the coefficients of (3), that $\phi_{m_1 \ldots m_k}(E_1, \cdots, E_k)$ is invariant under any simultaneous permutation of the m_j and the E_j, and that the ϕ which has no subscript ($k = 0$) is 1. It also is clear from (7) that (4) can be written in the form

$$(8) \qquad \sum_{m_1=0}^{\infty} \cdots \sum_{m_k=0}^{\infty} \phi_{m_1 \ldots m_k}(E_1, \cdots, E_k) = 1;$$

that (5) is equivalent to

$$(9) \qquad \sum_{m_k=0}^{\infty} \phi_{m_1 \ldots m_k}(E_1, \cdots, E_k) = \phi_{m_1 \ldots m_{k-1}}(E_1, \cdots, E_{k-1});$$

finally that, according to (6),

$$(10) \qquad \sum_{h=0}^{\infty} \phi_{h\, m_1 - h\, m_2 \ldots m_k}(E_0, E_1, E_2, \cdots, E_k) = \phi_{m_1 \ldots m_k}(E_0 + E_1, E_2, \cdots, E_k)$$

holds for any pair of disjoint sets E_0, E_1 of $\mathcal{E}$. In particular, if $k = 1$,

$$(10 \text{ bis}) \qquad \sum_{h=0}^{m} \phi_{h\, m-h}(E_0, E_1) = \phi_m(E_0 + E_1), \quad \text{where} \quad E_0 E_1 = O,$$

O denoting the empty set.

In addition,

$$(11) \qquad \Phi_{n_1 \ldots n_k}(E_1, \cdots, E_k) = \sum_{m_1=0}^{\infty} \cdots \sum_{m_k=0}^{\infty} \frac{(m_1 + n_1)! \cdots (m_k + n_k)!}{m_1! \cdots m_k!}$$

$$\text{times } \phi_{n_1 + m_1 \ldots n_k + m_k}(E_1, \cdots, E_k),$$

as is seen by comparing the partial derivatives of (3) and (7) at $(z_1, \cdots, z_k) = (1, \cdots, 1)$. If the same is done at $(z_1, \cdots, z_k) = (0, \cdots, 0)$, the result is

$$(12)\qquad \phi_{m_1\ldots m_k}(E_1,\cdots,E_k)=\frac{1}{m_1!\cdots m_k!}\sum_{n_1=0}^{\infty}\cdots\sum_{n_k=0}^{\infty}\frac{(-1)^{n_1+\ldots+n_k}}{n_1!\cdots n_k!}$$

$$\text{times}\quad \Phi_{m_1+n_1\ldots m_k+n_k}(E_1,\cdots,E_k).$$

The infinite (and, in general, non-recursive) linear substitutions (11), (12) are reciprocal (it being understood that the multiple series involved are absolutely convergent in virtue of (2)). According to (11) and (1 bis),

$$(11\text{ bis})\qquad \Psi_k(E_1,\cdots,E_k)=\sum_{n_1=0}^{\infty}\cdots\sum_{n_k=0}^{\infty}(m_1+1)\cdots(m_k+1)$$

$$\text{times}\quad \phi_{m_1+1\ldots m_k+1}(E_1,\cdots,E_k).$$

3. It is clear from (1) and (1 bis) that

$$(13)\qquad \Phi_{n_1\ldots n_k}\geqq 0$$

is equivalent to

$$(13\text{ bis})\qquad \Psi_n\geqq 0.$$

The assumption (13 bis), made in **1**, has not been used thus far. Actually, what will be needed is that

$$(14)\qquad \phi_{m_1\ldots m_k}\geqq 0.$$

It is obvious from (11) and (11 bis) that (14) is sufficient but not necessary for (13 bis), that is, for (13).

Suppose that not only (13) but also (14) is satisfied. Then (8) shows that, if k and the disjoint sets $E_1,\cdots,E_k$ of $\mathcal{E}$ are arbitrarily fixed, the k-fold sequences $\phi_{m_1\ldots m_k}(E_1,\cdots,E_k)$, where $m_j=0,1,2,\cdots;\ j=1,\cdots,k$, can be interpreted as representing a discrete probability distribution of events which are characterized by the variable subscript $(m_1,\cdots,m_k)$. More specifically, this probability distribution will be thought of as realized in terms of a repartition of "points" P on the sets E of $\mathcal{E}$. In other words, $\phi_{m_1\ldots m_k}(E_1,\cdots,E_k)$ will be interpreted as the probability that there be exactly m_j points P in E_j for all k values of j simultaneously, where the k set E_j of $\mathcal{E}$ can not overlap. This interpretation is consistent for mutually k disjoint sets E_j, since

(i) as has been seen after (7), the value $\phi_{m_1\ldots m_k}(E_1,\cdots,E_k)$ is invariant under any simultaneous permutation of the m_j and the E_j;

(ii) according to (9), the probability that $k-1$ of k specified sets contain specified numbers of points and the k-th set contain another specified number of points, is such as to supply, when summed over all possible values $(=0,1,2,\cdots)$ of this last specified number, precisely the probability that the $k-1$ sets contain the specified number of points;

(iii) according to (10), the probabilities of disjoint sets E_0, E_1 are logically additive for arbitrarily fixed $E_2, \cdots, E_k$.

According to (11), the expected value of the number of arrays consisting of an ordered n_1-uple of points P in $E_1, \cdots$, of an ordered n_k-uple of points P in E_k is $\Phi_{n_1 \ldots n_k}(E_1, \cdots, E_k)$. This is also clear from (1 bis), since (11 bis) shows that $\Psi_n(E_1, \cdots, E_n)$ represents the expected value of the number of ordered n-uples of points P such that the j-th point is in E_j for all n values of j simultaneously. It is understood that the E_j cannot overlap in these interpretations of the Φ and Ψ.

It is now clear why (13) or, equivalently, (13 bis) is necessary but not sufficient for (14). In fact, the expected values of the "distribution" which represent average values of non-negative integers can be non-negative if some of the "probabilities" are negative.

4. Suppose that the logical sum, S, of all sets E contained in the field $\mathcal{E}$ is a set E contained in $\mathcal{E}$, and that the whole space S (which can, but need not, be Euclidean) contains only a finite number of points P. This means that there exists a positive integer, ν, such that

$$(15_1) \quad \Phi_{n_1 \ldots n_k}(E_{n_1}, \cdots, E_{n_k}) \equiv 0 \text{ for every } n > \nu, \text{ where } n = n_1 + \cdots + n_k.$$

Suppose further that

$$(15_2) \quad \Phi_{n_1 \ldots n_k}(E_{n_1}, \cdots, E_{n_k}) = \frac{1}{(\nu - n)!} \Phi_{n_1 \ldots n_k\, \nu-n}(E_{n_1}, \cdots, E_{n_k}, E)$$

for every $n \leqq \nu$, where $n = n_1 + \cdots + n_k$, $E = S - (E_1 + \cdots + E_k)$.

This is a condition of consistency, if the number of points P in S is exactly ν.

It will be shown that, in this case of a finite number of points P in S, the requirement (14) is automatically satisfied in virtue of (13). In fact, it will be shown that, if (15_1) and (15_2) are satisfied, then

$$(16_1) \quad \phi_{n_1 \ldots n_k}(E_1, \cdots, E_k) \equiv 0 \text{ if } n > \nu, \text{ where } n = n_1 + \cdots + n_k,$$

while

$$(16_2) \quad \phi_{n_1 \ldots n_k}(E_1, \cdots, E_k) = \frac{1}{n_1! \cdots n_k!(\nu - n)!} \Phi_{n_1 \ldots n_k\, \nu-n}(E_1, \cdots, E_k, E)$$

if $n \leqq \nu$, where $n = n_1 + \cdots + n_k$,

$$E = S - (E_1 + \cdots + E_k); \; E_i E_j = O(i \neq j),$$

O denoting the empty set.

First, (16_1) is clear from (15_1) and (12). In order to verify (16_2), note that, under the assumptions of (16_2),

$$S^\nu = E_1^{n_1} \times \cdots \times E_k^{n_k} \times E^{\nu-n}$$

$$= \sum_{m_1+\ldots+m_k+m_{k+1}=\nu-n} \frac{(\nu-n)!}{m_1! \cdots m_k!\, m_{k+1}!} \text{ permutations of}$$

$$E_1^{m_1+n_1} \times \cdots \times E_k^{m_k+n_k} \times E^{m_{k+1}}$$

is a logical identity, where A^l denotes $A^{l-1} \times A$ or A according as $l > 1$ or $l = 1$, while $A^0 \times B = B$. Since the set functions Ψ_n are additive, it follows that

$$\Phi_{n_1\ldots n_k}(E_1, \cdots, E_k) = \sum_{m_1+\ldots+m_k+m_{k+1}=\nu-n} \frac{1}{m_1! \cdots m_k!\, m_{k+1}!}$$

$$\text{times } \Phi_{m_1+n_1\ldots m_k+n_k\, m_{k+1}}(E_1, \cdots, E_k, E)$$

in virtue of (1) and the assumptions (15_1), (15_2). This, when compared with the uniqueness of the reciprocal relations (11), (12), completes the proof of (16_2).

Note that, while the uniqueness of the correspondences expressed by the relations (11), (12) presupposes a restrictive condition of the type (2), this condition is trivially satisfied in the present case, (15_1). Thus (14) is a consequence of (13) in the case of (15_1), (15_2).

5. The sufficient condition of **4** for (14) implies a sufficient condition for (14) even if the number of points P in the space S is infinite. In fact, suppose that there are assigned, for every ν, functions $\Phi^{(\nu)}{}_{n_1\ldots n_k}$ such that (15_1) and (15_2) are satisfied by $\Phi = \Phi^{(\nu)}$, and let $\phi^{(\nu)}{}_{n_1\ldots n_k}$ denote the corresponding functions $\phi_{n_1\ldots n_k}$. Thus $\phi^{(\nu)}{}_{n_1\ldots n_k} \geqq 0$, by (16_1)-(16_2). It follows that, if the functions Φ are representable as averages,

$$\Phi_{n_1\ldots n_k}(E_1, \cdots, E_n) = \sum_{\nu=0}^{\infty} \lambda_\nu \Phi^{(\nu)}{}_{n_1\ldots n_k}(E_1, \cdots, E_k),$$

of the respective functions $\Phi^{(\nu)}$, where the weight factors λ_ν depend only on ν and satisfy

$$\lambda_\nu \geqq 0 \qquad \text{and} \qquad \sum_{\nu=0}^{\infty} \lambda_\nu = 1,$$

then the $\phi_{n_1\ldots n_k}$ which belong to the $\Phi_{n_1\ldots n_k}$ satisfy (14). In fact, it is easily verified that

$$\phi_{n_1\ldots n_k}(E_1, \cdots, E_k) = \sum_{\nu=0}^{\infty} \lambda_\nu \phi^{(\nu)}{}_{n_1\ldots n_k}(E_1, \cdots, E_k).$$

6. A rather particular case of a system of functions $\Phi_{n_1 \ldots n_k}$ which is of quite another type than the case considered in **4** is represented by the assumption of quasi-independence and corresponds, therefore, to an infinite chain:

$$\Phi_{n_1 \ldots n_k}(E_1, \cdots, E_k) = \prod_{j=1}^{k} \Phi_{n_j}(E_j).$$

It is understood that every Φ on the right has a single subscript; so that all $\phi_{n_1 \ldots n_k}(E_1, \cdots, E_k)$ are determined by the $\Phi_n(E)$ alone. For a fundamental characterization of this case of quasi-independence, cf. the end of **13.**

It turns out that, in the present case, (14) may but need not be satisfied.

First, it is easily verified from (12) that

$$\phi_{m_1 \ldots m_k}(E_1, \cdots, E_k) = \prod_{j=1}^{k} \phi_{m_j}(E_j);$$

so that (14) is equivalent to $\phi_m \geqq 0$, where m is arbitrary. Furthermore, (8) is equivalent to

$$\sum_{m=0}^{\infty} \phi_m(E) = 1,$$

while (10 bis) reduces to the addition rule

$$\sum_{h=0}^{m} \phi_h(E_0)\phi_{m-h}(E_1) = \phi_m(E_0 + E_1), \text{ where } E_0E_1 = O.$$

Since (12) implies that

$$\phi_m(E) = \frac{1}{m!} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \Phi_{n+m}(E),$$

it is clear that the present formulation, $\phi_m(E) \geqq 0$, of (14) may, but need not, be satisfied, if the present formulation, $\Phi_n(E) \geqq 0$, of (13) is satisfied. It also is seen from the last series that a sufficient condition for $\phi_m(E) \geqq 0$ consists in

$$\Phi_0(E) \geqq \Phi_1(E) \geqq \Phi_2(E) \geqq \cdots (\geqq 0),$$

where m and E are arbitrary.

Actually, it is sufficient to assume that every E be decomposable into a finite number of mutually disjoint subsets E in such a way that this monotony condition is satisfied for each of these subsets. In fact, $\phi_m(E) \geqq 0$ then holds for every m and for each of these subsets E and so, by the addition rule given before, for every m and for an arbitrary set E.

6 bis. Consider, finally, the case where the assumption of **6** concerning quasi-independence is replaced by the assumption of complete independence; so that $\Phi_n(E) = [\Phi_1(E)]^n$ and

$$\Phi_{n_1 \cdots n_k}(E_1, \cdots, E_k) = \prod_{j=1}^{k} [\Phi_1(E_j)]^{n_j}.$$

Then, according to (12),

$$\phi_{m_1 \cdots m_k}(E_1, \cdots, E_k) = \prod_{j=1}^{k} \left(\frac{[\Phi_1(E_j)]^{m_j}}{m_j!} \sum_{n=0}^{\infty} \frac{[-\Phi_1(E_j)]^n}{n!} \right) \equiv \prod_{j=1}^{k} \phi_{m_j}(E_j).$$

Thus

$$\phi_m(E) = \frac{a^m e^{-a}}{m!}, \text{ where } a = \Phi_1(E), \qquad (m = 0, 1, \cdots);$$

so that the distribution (belonging to $k = 1$) is the Poisson distribution of variance $a = \Phi_1(E)$.

In particular, (14) is implied by the assumption of complete independence.

Notice that the monotony condition, mentioned in **6**, reduces to

$$1 \geqq \Phi_1(E) \geqq [\Phi_1(E)]^2 \geqq [\Phi_1(E)]^3 \geqq \cdots$$

and is, therefore, satisfied only if E is so "small" that $\Phi_1(E) \leqq 1$.

PART II

The Space of Contingencies and its Lebesgue Measure

7. In what follows, the assumptions will be that the functions (1) satisfy (2) and (14), and that the field $\mathcal{E}$ of sets E has a finite or enumerable basis (cf. the beginning of **1**). This basis will be thought of as fixed. A set E which occurs in this basis will be denoted by E^*.

By a fundamental contingency, ω, will be meant a consistent allotment, assigning to every E^* a non-negative integer $m = m(E^*)$ which can be interpreted as follows: The set E^* of the basis of $\mathcal{E}$ contains exactly $m(E^*)$ points P. By the consistency of the allotment is meant, of course, that $m(E^*) = m(E^*_1) + m(E^*_2)$ whenever E^*_1 and E^*_2 are disjoint and $E^*_1 + E^*_2 = E^*$.

By a fundamental contingency set will be meant, of course, a set of fundamental contingencies ω; the ω which constitute the set being thought of as "points" of the set. It is understood that the basis of $\mathcal{E}$ which occurs in the assignment $m = m(E^*)$ of one ω is chosen to be the same for all points ω of the set of ω's.

By a primitive contingency set of order k will be meant a set consisting of all those fundamental contingencies for which $E^*_1, \cdots, E^*_k$ contain $m_1, \cdots, m_k$ points respectively, where the k non-negative integers m_j and the k sets E^*_j in the (fixed) basis of the field $\mathcal{E}$ are arbitrarily fixed but disjoint.

Consider the Borel field generated by the collection of all primitive contingency sets belonging to the fixed basis of $\mathcal{E}$. Let Γ denote an arbitrary set in this field, and Ω the particular Γ which is the logical sum of all sets Γ.

If Γ is a primitive contingency set, determined by k; $m_1, \cdots, m_k$ and $E^*_1, \cdots, E^*_k$, put

$$\mu(\Gamma) = \phi_{m_1 \ldots m_k}(E^*_1, \cdots, E^*_k) .$$

It is then clear from (8), (9), (10) and (14), that the requirement of complete additivity extends this function μ of the primitive contingency sets Γ to a unique Lebesgue measure $\mu = \mu(\Gamma)$, defined for all sets contained in the Borel field generated by the primitive contingency set Γ. Obviously,

$$\mu(\Omega) = 1.$$

In particular, Birkhoff's ergodic theorem is applicable on Ω, if there is given on the Borel sets Γ of Ω a μ-preserving transformation of Ω into itself.

7 bis. Let $\mathcal{E}^n$ and $\Psi_n(E_1, \cdots, E_n)$ be defined as at the beginning of **1**. Let $E^{(n)}$ denote a set in the Borel field generated by $\mathcal{E}^n$, and let $\Psi^{(n)}(E^{(n)})$ be the completely additive set function which reduces to $\Psi_n(E_1, \cdots, E_n)$ if $E^{(n)} = E_1 \times \cdots \times E_n$. It is seen from the definitions of **7**, that such an $E^{(n)}$ can be thought of as representing a set Γ.

In order to illustrate this remark, suppose first that there exists a ν for which the assumptions (15_1), (15_2) of **4** are satisfied. Then, according to (16_1), (16_2), (1) and the definition of μ in **7**,

$$\mu(\Gamma) = \frac{1}{\nu!} \Psi^{(\nu)}(\Gamma)$$

for every Γ of the type just mentioned. Furthermore, the assumption (14) of **7** is now automatically satisfied, by **4**.

It follows that, under the more general assumptions of **5**,

$$\mu(\Gamma) = \sum_{\nu=0}^{\infty} \frac{\lambda_\nu}{\nu!} \Psi^{(\nu)}(\Gamma^{(\nu)}),$$

where $\Gamma^{(\nu)}$ denotes the projection of Γ on the space $\Omega^{(\nu)}$ of exactly ν points P. Again, (14) is automatically satisfied, by **5**.

Still more general assumptions can be obtained by considering limits of systems of the type described in **5**. However, it remains problematic, in what sense, if any, is such an approximation possible in case of a given μ on a given Ω.

It is now seen that the measure μ introduced in **7** serves the purpose of supplying a substitute for the explicitly given measure of a system of the

elementary type, considered in **4** or **5**, in the case of a general system; a case in which the measure μ must be constructed, instead of being available *a priori*. Of course, (14) is not a consequence of (13) in this general case.

8. Suppose, for simplicity, that the sets E of the fixed $\mathfrak{E}$ are situated in a Euclidean space, S, of given dimension number, d; so that every E is a set of points x, where x is a vector with d components. Let $\mathfrak{E}$ be the field of all Borel sets E of S. As an enumerable basis of $\mathfrak{E}$, choose the collection of all those parallelepipeds which are parallel to the coördinate axes and have rational numbers for all d coördinates of each of their 2^d corners. In accordance with the notations of **7**, E^* will denote one of these rational parallelepipeds.

Thus, a basis contingency is given by a consistent allotment to every E^* of a non-negative integer $m = m(E^*)$ which can be interpreted as the number of points P contained in E^*. Since the allotment is consistent, $m(E^*) = m(E^*_1) + m(E^*_2) + \cdots$ for every decomposition of a basis parallelepiped E^* into mutually disjoint basis parallelepipeds $E^*_1, E^*_2, \cdots$.

For a fixed E^*, consider a sequence of such decompositions, having the property that the $(l+1)$-th decomposition is obtained by a subdivision of the l-th and the maximum diameter of the parallelepipeds occurring in the l-th decomposition tends to 0 as $l \to \infty$. For every fixed l, not more than $m(E^*)$ of the parallelepipeds of the l-th decomposition contain at least one point P. Since $m(E^*)$ is independent of l, it follows that in the sequence of successive decompositions of E^* not more than $m(E^*)$ nested sequences of parallelepipeds contain points P. Let $\bar{m}(E^*)$ denote the number $(\geqq 0)$ of these non-empty nested sequences of parallelepipeds. Thus $\bar{m}(E^*) \leqq m(E^*)$, where $\bar{m}(E^*) < m(E^*)$ cannot be excluded, since some of the points P can be multiple. On the other hand, $\bar{m}(E^*) \geqq 1$ unless $m(E^*) = 0$.

9. If x is one of the $\bar{m}(E^*)$ points of E^* to which the $\bar{m}(E^*)$ nested sequences of parallelepipeds tend as $l \to \infty$, there exists a non-negative integer $i = i(x)$ such that the l-th parallelepiped in the sequence of nested parallelepipeds about x contains exactly $i(x)$ points P for every sufficiently large x. Furthermore, the sum of the $\bar{m}(E^*)$ integers $i(x)$ is $m(E^*)$, if suitable precaution is taken for the faces of E^*, i. e., for the boundary of the set E^*. (It is possible to take such a precaution, since E^* can be chosen arbitrarily in the basis of the Euclidean space S.) If x is any point which lies in the interior of E^* and is distinct from the $\bar{m}(E^*)$ points x for which $i(x)$ was just defined, put $i(x) = 0$.

Since E^* can be chosen arbitrarily, there is now defined a non-negative integer $i = i(x)$ as a function of the position x on the Euclidean space S.

Since there exist in every E^* at most a finite number ($\leqq m(E^*)$, possibly 0) of points x for which $i(x) \neq 0$, and since the set of all E^* is enumerable, there exists in S an at most enumerable set of distinct points X which have no finite cluster point in S and are characterized by the fact that the index $i(x)$ of a point x of S is zero or a positive integer according as x is or is not an X. Let $\{X\}$ denote the sequence of all points X of S. It is understood that the set $\{X\}$ can be finite.

10. It is clear from **8** that a basis contingency is equivalent to a consistent allotment of a non-negative integer $\bar{m}(E^*)$ and of $\bar{m}(E^*)$ points of E^*, along with the multiplicities of these points, for each of the parallelepipeds E^* which form a basis of the field $\mathfrak{E}$ of the Borel sets E of S. It follows, therefore, from **9**, that a basis contingency is equivalent to the assignment of an arbitrary sequence $\{X\}$ of distinct points X of S which have no finite cluster point, and of an arbitrary positive integer $i(X)$ as a function of the position X on $\{X\}$. In fact, a basis contingency then is defined by assigning that every basis set E^* contain exactly $m(E^*) = \Sigma\, i(X)$ points P, where the summation runs through those points X of $\{X\}$ which are in E^*.

While only *fundamental contingencies,* and not contingencies (unqualified) have been defined so far, it is now possible to define *contingencies,* as follows: A contingency is given by a consistent allotment to every bounded Borel set E in S of a non-negative integer, $m(E)$, which represents the number of points P contained in E and is extended, either as a non-negative integer $m(E)$ or as $m(E) = \infty$, to every (not necessarily bounded) Borel set E of S by the requirement of complete additivity. In fact, this definition of a contingency is equivalent to the assignment of an arbitrary positive integer $i(X)$ as a function of the position X on an arbitrary sequence $\{X\}$ of points of the Cartesian space S which have no finite cluster point. It is understood that the number ($\geqq 0, \leqq \infty$) of points P contained in an arbitrary Borel set E of S is then given by $m(E) = \Sigma\, i(X)$, where the summation runs through those points X of $\{X\}$ which are in E.

11. In **7,** the symbol ω was used to denote an arbitrary fundamental contingency. Let ω now denote an arbitrary contingency. While the sets which in **7** were denoted by Γ were there introduced by a Borel extension of the collection of all primitive contingency sets (instead of as sets consisting of contingencies as points), it is now possible to consider every set Γ as a Borel set of points ω. In particular, the set of all points ω is the particular Γ which in **7** was denoted by Ω. Thus Ω is now a space consisting of all contingencies ω as points, and $\mu(\Gamma)$ is a measure, defined on the field of sets Γ which are subsets of the contingency space Ω.

The definition of $\bar{m}(E^*)$ in **8** is extended from basis sets E^* to arbitrary Borel sets E of S, if $\bar{m}(E)$ denotes the number of those points X of $\{X\}$ which are in E. Since $i(X)$ is a positive integer, $\bar{m}(E) \leqq m(E)$. The sign of equality holds for every E if and only if $i(X) = 1$ for every X. Since if x is an arbitrary point of S, then $i(x) = 0$ or $i(x) \geqq 1$ according as x is or is not an X, it is clear that $\bar{m}(E) = m(E)$ for every E if and only if $i(x)$ is either 0 or 1 for every x.

It follows that, if $\bar{m}(E) = m(E)$ for every E, then a contingency is equivalent to an arbitrary allotment of a sequence $\{X\}$ of points X of S which do not have a finite cluster point. In fact, the allotment of the function $i(X)$ on $\{X\}$, as required by **10**, is then given by $i(X) \equiv 1$.

In what follows, there will be delimited cases in which $\bar{m}(E) \equiv m(E)$, i. e. $i(X) \equiv 1$, is true, though not necessarily for all, at least for almost all, points ω of the contingency space Ω. By this is meant that, in the cases to be considered, the exceptional contingencies are of measure 0 with respect to the μ-measure on Ω (cf. **7** and **10**). Since this measure $\mu(\Gamma)$ on Ω was defined (**7**) in terms of the system of set functions $\phi_{m_1 \ldots m_k}(E_1, \cdots, E_k)$, it depends, of course, on the choice of this system of set functions whether $i(X) \equiv 1$ is or is not true for almost all contingencies on Ω; so that what must be delimited is a class of suitable systems of set functions $\phi_{m_1 \ldots m_k}$.

12. In view of (11), (12), (1) and the assumption (1), it will be sufficient to consider classes of suitable systems of set functions $\Psi_n(E_1, \cdots, E_n)$. Actually, it turns out that, in order to obtain a useful sufficient criterion for the validity of $i(X) \equiv 1$ almost everywhere on Ω, it is not necessary to consider the full sequence $\Psi_n(E_1, \cdots, E_n)$, where $n = 1, 2, \cdots$, but only Ψ_1 and Ψ_2.

In fact, suppose that there exists for every point x of the Euclidean space S and for every $\epsilon > 0$ a parallelepiped $E^*_x(\epsilon)$ about x in such a way that

$$(*) \quad \Psi_2(E, E) \leqq \epsilon \Psi_1(E) \text{ whenever } E \text{ is in } E^*_x(\epsilon)$$

(it being understood that E denotes a Borel set in S). It will be shown that, if this condition is satisfied, the set of those points ω of the contingency space Ω for which $0 \leqq i(x) \leqq 1$ does not hold for all points x of S is of μ-measure 0.

Let E^* be a fixed parallelepiped in the enumerable basis of S (cf. **8**). Let $\bar{m}(E^*)$ denote the same integer as in **8**. It is then clear from **9-10** that the set of those fundamental contingencies for which E^* contains exactly $\bar{m}(E^*)$ distinct points P is a primitive contingency set in the sense of **7**. Consider these point groups as *ordered* point groups, and let E^* contain the point P.

For every point x of the parallelepiped E^* and for a given $\epsilon > 0$, choose the parallelepiped $E^*_x(\epsilon)$ in accordance with (*). Then, if E^* is closed and $\epsilon > 0$ is fixed, the Heine-Borel theorem assures that there is a finite number of points x in E^* such that the $E^*_x(\epsilon)$ which belong to these x cover the whole of E^*. Hence, there exists a finite number of mutually disjoint parallelepipeds E^*_l such that $\sum_l E^*_l = E^*$ and, for every l,

$$\Phi_2(E^*_l) = \Psi_2(E^*_l, E^*_l) \leqq \epsilon \Psi_1(E^*_l).$$

Thus

$$\sum_l \Phi_2(E^*_l) = \sum_l \Psi_2(E^*_l, E^*_l) \leqq \epsilon \Psi_1(E^*),$$

since $\Psi_1(E)$ is additive (**1**). In view of the interpretation of Φ_n in **3**, the sum on the left of the last inequality cannot be less than the expected value of the number of those points X in E at which $i(X) > 1$. It follows, therefore, from the definition of μ in **7**, that, since $\epsilon > 0$ is arbitrarily small, the set of those contingencies for which the *first* of the ordered group of $\bar{m}(E^*)$ points X in E^* satisfies $i(X) > 1$ is of μ-measure zero. Since the sum of a sequence of contingency sets of μ-measure 0 is of μ-measure 0, the proof is complete.

13. Suppose that

(i) $$\Psi_k(E_1, E_2, \cdots, E_k) \leqq c^k \Psi_1(E_1)\Psi_1(E_2) \cdots \Psi_1(E_k)$$

holds for a sufficiently large constant c, and that

(ii) $$\Psi_1(E) \to 0 \text{ as } E \to x$$

(that is, that $\Psi_1(E)$ tends to 0 whenever a nested sequence of Borel sets E shrinks to any fixed point x of the Euclidean space S). It is clear that (ii) implies, in virtue of (i), the restriction (*) of **12**. Furthermore, (2) can be replaced by the sharper estimate

(i bis) $$\Phi_{n_1 \ldots n_k}(E_1, \cdots, E_k) \leqq c^{n_1+\ldots+n_k} [\Phi_1(E_1)]^{n_1} [\Phi_1(E_2)]^{n_2} \cdots [\Phi_1(E_k)]^{n_k},$$

which is obvious from (1) by the assumption (i).

Since Φ_1 is additive and non-negative (**1**), it is clear that $\Phi_1(E)$ can be thought of as defining a Lebesgue measure on the d-dimensional Euclidean space S. Then $\Phi_1(E_1)\Phi_1(E_2) \cdots \Phi_1(E_k)$ defines a product measure on the product space $S^k = S^{k-1} \times S$, where $S^1 = S$.

It is easily seen from (i) and (ii) that the set function defined by the Lebesgue extension of the set function (1 bis), **1** on S^k is absolutely continuous with respect to the product measure define by the corresponding extension of

$\Phi_1(E_1)\Phi_1(E_2)\cdots\Phi_1(E_k)$. Accordingly, there exists on the product space S^k a function δ_k of the position $(x_1,\cdots,x_k)$ such that

$$\Psi_k(E_1,\cdots,E_k)=\int_{E_1}\cdots\int_{E_k}\delta_k(x_1,\cdots,x_k)\,d_{x_1}\Phi_1(x_1)\cdots d_{x_k}\Phi_1(x_k),$$

where every x_j represents a point of the Euclidean space $S^1=S$ of given dimension number d. Needless to say, the L-integrable point function δ_k is determined by the set function Ψ_k almost everywhere (that is, up to an $(x_1,\cdots,x_k)$-set of vanishing product measure), and is non-negative (almost everywhere), since $\Psi_k\geqq 0$ (**1**). Thus, from (i),

$$0\leqq\delta_k(x_1,\cdots,x_k)\leqq c^k$$

(almost everywhere)

It has been seen in **3** that $\Psi_k(E_1,\cdots,E_k)$ represents the expected value of the number of ordered k-uples of points the j-th of which is in E_j for all k values of j simultaneously. It follows, therefore, from the result of **12** and from the preceding integral representation of Ψ_k, that $\delta_k(x_1,\cdots,x_k)$ is the density of probability of the ordered k-uples at $(x_1,\cdots,x_k)$.

Obviously, the case of quasi-independence (**6**) results if each of the densities (with respect to product measure!), $\delta_k(x_1,\cdots,x_k)$, is independent of the position $(x_1,\cdots,x_k)$ on S^k.

PART III

Functions of the Contingencies

14. In what follows, the assumption will be that the conditions of **7**, **8** and **12** (but not necessarily those of **13**) are satisfied. In particular, a contingency (unqualified) is defined, by **11**, as a point ω of space Ω which carries a Lebesgue measure μ such that $\mu(\Omega)=1$.

For k given Borel sets E_j of the underlying d-dimensional Euclidean space S and for k given non-negative integers m_j, define a function

$$\phi=\phi(\omega)\equiv\phi^{\omega}{}_{m_1\ldots m_k}(E_1,\cdots,E_k)$$

of the position ω on Ω as follows: The point function $\phi^{\omega}{}_{m_1\ldots m_k}(E_1,\cdots,E_k)$ of ω is the characteristic function of that set

$$\Gamma=\Gamma_{m_1\ldots m_k}(E_1,\cdots,E_k)$$

of contingencies in which E_j contains exactly m_j points P. Since this $\Gamma_{m_1\ldots m_k}(E_1,\cdots,E_k)$ obviously is a μ-measurable subset of Ω, its characteristic function is integrable over Ω. Furthermore, if the k subscripts of this

characteristic function are varied, a straightforward counting verifies the orthogonality relation

$$(*)\quad \int_{\Omega} \phi^{\omega}{}_{m_1\cdots m_k}(E_1,\cdots,E_k)\phi^{\omega}{}_{l_1\cdots l_k}(E_1,\cdots,E_k)\mu(\delta\Omega)$$
$$= \begin{cases} 0 \text{ if } (m_1,\cdots,m_k) \neq (l_1,\cdots,l_k), \\ \phi_{m_1\cdots m_k}(E_1,\cdots,E_k) \text{ if } (m_1,\cdots,m_k) = (l_1,\cdots,l_k), \end{cases}$$

where $E_1,\cdots,E_k$ are fixed and $\phi_{m_1\cdots m_k}(E_1,\cdots,E_k)$ denotes the same set function as in **2-3**.

15. Let a non-negative integer n and n Borel sets $E_1,\cdots,E_n$ of S be given. Define a non-negative integer

$$\Psi = \Psi(\omega) = \Psi_n{}^{\omega}(E_1,\cdots,E_n)$$

as a function of the position ω on Ω as follows: $\Psi_n{}^{\omega}(E_1,\cdots,E_n)$ is the number of ordered n-uples $(P_1\cdots P_n)$ which correspond to ω and to the restriction that the point P_j is in E_j for $j=1,\cdots,n$, where ω is a given contingency.

Corresponding to (1), define a function

$$\Phi = \Phi(\omega) \equiv \Phi^{\omega}{}_{n_1\cdots n_k}(E_1,\cdots,E_k)$$

of the position ω on Ω by placing

$$(\mathrm{I})\quad \Phi^{\omega}{}_{n_1\cdots n_k}(E_1,\cdots,E_k) = \Psi_n{}^{\omega}(E_1,\cdots,E_1,\cdots,E_k,\cdots,E_k), \quad \text{where } n = n_1+\cdots+n_k,$$

it being understood that every E_j occurs exactly n_j times in $\Psi_n{}^{\omega}$. In particular

$$(\mathrm{I\ bis})\quad \Psi_k{}^{\omega}(E_1,\cdots,E_k) = \Phi^{\omega}{}_{1\cdots 1}(E_1,\cdots,E_k), \text{ where } 1+\cdots+1=n.$$

It is easily seen from the definition of $\phi^{\omega}{}_{m_1\cdots m_k}(E_1,\cdots,E_k)$ in **14** that, corresponding to (11),

$$(\mathrm{II})\quad \Phi^{\omega}{}_{n_1\cdots n_k}(E_1,\cdots,E_k) = \sum_{m_1=0}^{\infty}\cdots\sum_{m_k=0}^{\infty} \frac{(m_1+n_1)!\cdots(m_k+n_k)!}{m_1!\cdots m_k!}$$
$$\text{times } \phi^{\omega}{}_{n_1+m_1\cdots n_k+m_k}(E_1,\cdots,E_k);$$

in fact, all but one of the terms of the non-negative series (II) vanish. Similarly, corresponding to (12),

$$(\mathrm{III})\quad \Phi^{\omega}{}_{m_1\cdots m_k}(E_1,\cdots,E_k) = \frac{1}{m_1!\cdots m_k!}\sum_{n_1=0}^{\infty}\cdots\sum_{n_k=0}^{\infty}\frac{(-1)^{n_1+\cdots+n_k}}{n_1!\cdots n_k!}$$
$$\text{times } \Phi^{\omega}{}_{m_1+n_1\cdots m_k+n_k}(E_1,\cdots,E_k);$$

it being clear that all but a finite number of the terms of the series (III) vanish.

16. While (II) and (III) are obvious identities for every fixed ω, it is by no means obvious that both of these expansions are valid in the (L^2)-mean of the μ-measure on Ω also. It will now be shown that such happens to be the case.

In other words, it will be shown that the quadratic mean error of the partial sums of the series, an error represented for (III) by

$$\text{(IV)} \quad \int_\Omega \Big[\phi^\omega_{m_1\dots m_k}(E_1,\cdots,E_k) - \frac{1}{m_1!\cdots m_k!}\sum_{n_1=0}^{N_1}\cdots\sum_{n_k=0}^{N_k}\frac{(-1)^{n_1+\dots+n_k}}{n_1!\cdots n_k!} \text{ times } \Phi^\omega_{m_1+n_1\dots m_k+n_k}(E_1,\cdots,E_k)\Big]^2\mu(d\Omega),$$

tends to zero as $N_1\to\infty,\cdots,N_k\to\infty$, if k and $E_1,\cdots,E_k$ are arbitrarily fixed; and that the corresponding relation holds for (II) also.

In view of the orthogonality relation, mentioned at the end of **14**, the validity of the expansions (II), (III) in the mean can be interpreted as expressing the completeness of the orthogonal system at hand. Actually, it is precisely this completeness property that will be needed in the sequel.

16 bis. The proof proceeds as follows:

Since the series (II) cannot have more than one non-vanishing term for a fixed ω, it is easily verified from the orthogonality relations, mentioned at the end of **14**, that

$$\int_\Omega \phi^\omega_{m_1\dots m_k}(E_1\cdots,E_k)\Phi^\omega_{n_1\dots n_k}(E_1,\cdots,E_k)\mu(d\Omega)$$

$$= \begin{cases} \dfrac{m_1!\cdots m_k!}{(m_1-n_1)!\cdots(m_k-n_k)!}\,\phi_{m_1\dots m_k}(E_1,\cdots,E_k) \text{ if } m_1>n_1,\cdots,m_k>n_k, \\ 0 \text{ otherwise} \end{cases}$$

and

$$\int_\Omega \Phi^\omega_{m_1\dots m_k}(E_1\cdots,E_k)\Phi^\omega_{n_1\dots n_k}(E_1,\cdots,E_k)\mu(d\Omega)$$

$$= \sum_{j_1=\min(m_1,n_1)}^{\infty}\cdots\sum_{j_k=\min(m_k,n_k)}^{\infty}\frac{(j_1!\cdots j_k!)^2\ \phi_{j_1\dots j_k}(E_1,\cdots,E_k)}{(j_1-n_1)!\cdots(j_k-n_k)!(j_1-m_1)!\cdots(j_k-m_k)!}.$$

Hence, the mean square error, represented by the integral (IV), is identical with

$$- \phi_{m_1 \ldots m_k}(E_1, \cdots, E_k) - \frac{2}{m_1! \cdots m_k!} \sum_{n_1=\min(m_1,N_1)}^{\infty} \cdots \sum_{n_k=\min(m_k,N_k)}^{\infty}$$

$$\frac{(-1)^{n_1+\ldots+n_k} m_1! \cdots m_k!}{n_1! \cdots n_k! (m_1 - n_1)! \cdots (m_k - n_k)!} \phi_{m_1 \ldots m_k}(E_1, \cdots, E_k)$$

$$+ \frac{1}{(m_1! \cdots m_k!)^2} \sum_{n_1=0}^{N_1} \cdots \sum_{n_k=0}^{N_k} \sum_{l_1=0}^{N_1} \cdots \sum_{l_k=0}^{N_k} \frac{(-1)^{n_1+\ldots+n_k+l_1+\ldots+l_k}}{n_1! \cdots n_k! l_1! \cdots l_k!} \sum_{j_1=m_1+\min(n_1,l_1)}^{\infty} \cdots \sum_{j_k=m_k+\min(n_k,l_k)}^{\infty}$$

$$\frac{(j_1! \cdots j_k!)^2 \, \phi_{j_1 \ldots j_k}(E_1, \cdots, E_k)}{(j_1 - m_1 - n_1)! \cdots (j_k - m_k - n_k)! (j_1 - m_1 - l_1)! \cdots (j_k - m_k - l_k)!},$$

and reduces therefore, if $N_1 \geqq m_1, \cdots, N_k \geqq m_k$, to

$$- \phi_{m_1 \ldots m_k}(E_1, \cdots, E_k) + \sum_{j_1=m_1}^{\infty} \cdots \sum_{j_k=m_k}^{\infty} \phi_{j_1 \ldots j_k}(E_1, \cdots, E_k) (j_1! \cdots j_k!)^2 \text{ times}$$

$$\sum_{n_1=\min(N_1,j_1-m_1)}^{\infty} \cdots \sum_{n_k=\min(N_k,j_k-m_k)}^{\infty} \sum_{l_1=\min(N_1,j_1-m_1)}^{\infty} \cdots \sum_{l_k=\min(N_k,j_k-m_k)}^{\infty}$$

$$\frac{(-1)^{n_1+\ldots+n_k+l_1+\ldots+l_k}}{n_1! \cdots n_k! (j_1 - m_1 - n_1)! \cdots (j_k - m_k - n_k)! (j_1 - m_1 - l_1)! \cdots (j_k - m_k - l_k)!}$$

or simply to

$$- \phi_{m_1 \ldots m_k}(E_1, \cdots, E_k) + \sum_{j_1=m_1}^{\infty} \cdots \sum_{j_k=m_k}^{\infty} \phi_{j_1 \ldots j_k}(E_1, \cdots, E_k) (j_1! \cdots j_k!)^2 \text{ times}$$

$$\left(\sum_{n_1=\min(N_1,j_1-m_1)}^{\infty} \cdots \sum_{n_k=\min(N_k,j_k-m_k)}^{\infty} \frac{(-1)^{n_1+\ldots+n_k}}{n_1! \cdots n_k! (j_1 - m_1 - n_1)! \cdots (j_k - m_k - n_k)!} \right)^2.$$

Since k and $m_1, \cdots, m_k$ are fixed, it follows that, as $N_1 \to \infty, \cdots, N_k \to \infty$, the integral (IV) is majorized by a constant multiple of

$$\sum_{j_1=N_1+m_1}^{\infty} \cdots \sum_{j_k=N_k+m_k}^{\infty} \phi_{j_1 \ldots j_k}(E_1, \cdots, E_k) \left(\frac{j_1! \cdots j_k! \, 2^{j_1+\ldots+j_k}}{(j_1 - m_1)! \cdots (j_k - m_k)!} \right)^2.$$

In order to estimate the coefficients of ϕ in this series, notice that, by (2), the function (3) is an entire function of the exponential type in the $z_j - 1$; hence, it is an entire function of the exponential type in the z_j also. This means, in view of (7), that

$$| \phi_{m_1 \ldots m_k}(E_1, \cdots, E_k) | < \frac{C^{m_1+\ldots+m_k}}{m_1! \cdots m_k!}$$

holds for a sufficiently large C which is independent of $(m_1, \cdots, m_k)$. Hence, the dominant series of (IV), found before, is majorized by a constant multiple of

$$\sum_{j_1=N_1+m_1}^{\infty} \cdots \sum_{j_k=N_k+m_k}^{\infty} \frac{C^{j_1+\ldots+j_k}}{j_1! \cdots j_k!} \left(\frac{j_1! \cdots j_k! \, 2^{j_1+\ldots+j_k}}{(j_1 - m_1)! \cdots (j_k - m_k)!} \right)^2.$$

This series can be written in the form

$$\sum_{j_1=N_1}^{\infty} \cdots \sum_{j_k=N_k}^{\infty} \frac{(4C)^{j_1+\cdots+j_k}}{j_1! \cdots j_k!} \frac{(j_1+m_1)! \cdots (j_k+m_k)!}{j_1! \cdots j_k!}$$

and tends, therefore, to zero as $N_1 \to \infty, \cdots, N_k \to \infty$, the integers k and $m_1, \cdots, m_k$ being fixed. This proves that (IV) tends to zero as $N_1 \to \infty, \cdots, N_k \to \infty$.

In other words, the expansion (III) is valid in the mean. The proof of the corresponding statement of **16** concerning the expansion (II) is similar.

17. According to **11**, a contingency ω can be thought of as consisting of certain assignments of definite points P in the Euclidean space S; while **12** assures that these points are all distinct. Thus, it is clear from the definition of $\Psi_n{}^{\omega}(E_1, \cdots, E_n)$ in **15** that, for every fixed n,

$$\Psi_n{}^{\omega}(E_1, \cdots, E_n) = \Sigma \cdots \Sigma \psi^{\omega}{}_{E_1 \ldots E_n}(P_1, \cdots, P_n),$$

where the n summation sign $\Sigma \cdots \Sigma$ represents summation over all n-uples of distinct points $P_1, \cdots, P_n$ corresponding to the contingency ω, and $\psi^{\omega}{}_{E_1 \ldots E_n}(P_1, \cdots, P_n)$ denotes the characteristic function of the product set $E_1 \times \cdots \times E_n$.

Correspondingly, the completeness theorem proved in **16** bis (cf. the end of **16**), when applied to the $\psi_n{}^{\omega}$ insteal of the $\Phi^{\omega}{}_{n_1 \ldots n_k}$, can be expressed by saying that the system

$$1, \quad \Sigma \psi_E{}^{\omega}(P), \quad \Sigma\Sigma \psi^{\omega}{}_{E_1 E_2}(P_1, P_2), \quad \Sigma\Sigma\Sigma \psi^{\omega}{}_{E_1 E_2 E_3}(P_1, P_2, P_3), \cdots$$

is a closed (L^2) system of functions on the contingency space. Since these functions are independent, they could be ortho-normalized.

18. The preceding completeness theory of contingencies is relevant for an ergodic theory of contingencies. This will now be illustrated by the simplest case, that of the translation group.

Suppose that the functions (1 bis), **1** are given so as to remain invariant if the sets $E_1, \cdots, E_k$ of the d-dimensional Euclidean space S are subject to the same translation, which is allowed to be arbitrary. In other words, let, for every k and for arbitrary $E_1, \cdots, E_k$,

$$\Psi_k(E_1 + s, \cdots, E_k + s) = \Psi_k(E_1, \cdots, E_k),$$

where s is an arbitrary vector with d components and $E + s$ denotes the set of all those vectorial sums $e + s$ for which the vector $x = e$ represents a point of the Borel set E of S. Using the assumptions and notations of **13**, one can also say that, on the one hand,

$$\delta_k(x_1 + s, \cdots, x_k + s) = \delta_k(x_1, \cdots, x_k),$$

and, on the other hand, $d_x\Psi_1$ is proportional to the d-dimensional Euclidean volume element, dx. Accordingly, the representation of $\Psi_k(E_1, \cdots, E_k)$ in **13** reduces to

$$\text{(i)} \qquad \Psi_k(E_1, \cdots, E_k) = \int_{E_1} \cdots \int_{E_k} \rho_k(x_1, \cdots, x_k)\, dx_1 \cdots dx_k,$$

where $\rho_k(x_1, \cdots, x_k)$ is proportional to the density $\delta_k(x_1, \cdots, x_k)$,

$$\rho_k(x_1, \cdots, x_k) = (\text{const.})^k\, \delta_k(x_1, \cdots, x_k), \qquad (\delta_0 \equiv 1);$$

so that

$$\text{(ii)} \qquad \rho_k(x_1 + s, \cdots, x_k + s) = \rho_k(x_1, \cdots, x_k)$$

and, according to **13**,

$$\text{(iii)} \qquad 0 \leqq \rho_k(x_1, \cdots, x_k) \leqq (\text{Const.})^k.$$

It is clear from **1-12** that addition of an arbitrary constant vector, s, to every point X of a contingency $\omega = \{X\}$ generates a μ-measure-preserving transformation, say τ_s, of the contingency space Ω into itself. All transformations τ_s together, which belong to translations s of the d-dimensional Euclidean space S, constitute a group. Clearly, the multi-dimensional ergodic theorem is applicable to this group. Accordingly, if $F(\omega)$ is a function of class (L) on Ω, and if the point ω of Ω is fixed not on a certain subset of μ-measure 0, the s-average of F $(\tau_s\omega)$ over the interior of an s-sphere about the unit of the group τ_s tends to a limit when this sphere increases indefinitely.

19. In view of the physical applications alluded to in the Introduction, it is of fundamental importance to have a criterion which assures that the "flow" τ_s on Ω makes almost all "paths" on the configurations statistically independent and such as to correspond to the case of asymptotic equidistribution.

It will be shown that a simple sufficient criterion to this effect is represented by the approximate independence of remote regions of S; that is, by the assumption that

$$(*) \qquad \rho_k(x_1, \cdots, x_l, x_{l+1} + s, \cdots, x_k + s) \to \rho_l(x_1, \cdots, x_l)\rho_{k-l}(x_{l+1}, \cdots, x_k)$$

as $|\, s \,| \to \infty$, where k and l are arbitrary. In fact, it will be shown that, if (*) is satisfied, τ_s is a mixture.

In particular, (*) is sufficient for the metrical transitivity of τ_s, that is, for the situation in which the s-average of $F(\tau_s\omega)$ is, for almost all points ω of Ω, independent of ω and, therefore equals

$$\int_{\Omega} F(\omega)\mu(d_{\omega}\Omega),$$

where $F(\omega)$ is any fixed function of class (L) on Ω.

19 bis. The proof proceeds as follows:

First, every $\rho_k(x_1, \cdots, x_k)$ is bounded, by (iii), **18**. Hence, it is clear from (i), **18** and (*), **19** that

$$\Psi_k(E_1, \cdots, E_l, E_{l+1} + s, \cdots, E_k + s) \to \Psi_l(E_1, \cdots, E_l)\Psi_{k-l}(E_{l+1}, \cdots, E_k)$$

as $|s| \to \infty$. It follows, therefore, from (2), **1** and from the connections between the set functions Ψ, Φ, ϕ, that

$$\phi_{m_1 \cdots m_k}(E_1, \cdots, E_l, E_{l+1} + s, \cdots, E_k + s) \to \phi_{m_1 \cdots m_l}(E_1, \cdots, E_l)\phi_{m_{l+1} \cdots m_k}(E_{l+1}, \cdots, E_k)$$

Cf. **1-3** and **16** bis.

Since a basis of the measure μ on Ω can be obtained by assigning the measure $\phi_{m_1 \cdots m_k}(E_1, \cdots, E_k)$ to that set of contingencies ω for which E_j contains exactly m_j points P for $j = 1, \cdots, k$, it is now seen from **17** that, if F is the characteristic function of any $\mu \times \mu$-measurable set on the product space $\Omega \times \Omega$, then

$$\int_{\Omega} F(\Omega, \tau_s\Omega)\mu(d_{\omega}\Omega) \to \int_{\Omega}\int_{\Omega} F(\omega, \bar{\omega})\mu(d_{\omega}\Omega)\mu(d_{\bar{\omega}}\Omega)$$

as $|s| \to \infty$. Thus τ_s is a mixture on Ω, i. e., the product flow, $\tau_s \times \tau_s$, is metrically transitive on $\Omega \times \Omega$.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY,
THE JOHNS HOPKINS UNIVERSITY.

GROUPS OF ALGEBRAS OVER AN ALGEBRAIC NUMBER FIELD.*

By S. MacLane and O. F. G. Schilling.

In this note we shall discuss some new aspects of the theory of normal simple algebras over an algebraic number field. The algebras considered are split by a finite normal extension which is the join of two normal subfields. Moreover, the ramifications of the algebras are to be prime to the discriminant of the given normal field. We propose to find conditions insuring that the group of algebras split by the top field be the join of the groups of algebras split by the given component fields. This query leads to certain curious facts concerning the associated Galois groups. One of our results states that the group of algebras for any abelian field is always the join of the groups of algebras split by any set of subfields whose join is the top field. In the general case we have succeeded in finding a partial description of the Galois groups in terms of the condition on the associated algebras. The method of proof of our theorems consists in a reduction of problems on algebras to problems on groups by means of the theory of invariants of algebras.

Let K be a finite normal extension with Galois group $\Gamma = \{\sigma, \cdots\}$ over a finite algebraic field F. Suppose that the field K/F is given as the join $K' \bigcup K''$ of two normal subfields K'/F and K''/F with the respective Galois groups $\Gamma' = \{\sigma', \cdots\}$ and $\Gamma'' = \{\sigma'', \cdots\}$. Let $\Delta = \{\delta, \cdots\}$ be the Galois group of the intersection $(K' \bigcap K'')/F$. Let T' and T'' be the homomorphisms mapping the groups Γ' and Γ'' upon Δ. Then [1] the Galois group Γ can be described as the group of all pairs (σ', σ'') such that $\sigma'\mathrm{T}' = \sigma''\mathrm{T}''$. We call Γ the *subdirect product* of the factor groups Γ', Γ''.

In the sequel we are necessarily led to consider a class of fields, in which each field K is the join of two normal subfields K' and K'' such that the following hypothesis (H) holds:

H: For every prime l and positive integer m such that l^m is the order of an element τ' in Γ' and of an element τ'' in Γ'' there exists at least one element $\sigma = (\sigma', \sigma'')$ in Γ whose components σ', σ'' have orders divisible by l^m.

Let M be the product of all those finite and infinite prime divisors of F

* Received December 6, 1941; Presented to the Society, December 29-31, 1941.

[1] For an outline of the proof for the following assertion see S. MacLane and O. F. G. Schilling, "A formula for the direct product of crossed product algebras," *Bulletin of the American Mathematical Society*, vol. 48 (1942), pp. 108-114.

which are ramified in K. We shall consider the groups of algebra classes S, S', S'' which are prime to M and split by the normal fields K, K', K'', respectively.[2] The group S contains as subgroups both S', S'' and their join $S' \cup S''$. We prove

THEOREM 1. *Let K, K', K'' be a triple of normal fields with $K = K' \cup K''$. Then $S = S' \cup S''$ if and only if hypothesis (H) holds.*

Proof. We first prove that the condition is sufficient. It is enough to prove the assertion for algebras of prime power degree l^r. Such algebras we shall term l-primary algebras. For any algebra S is similar[3] to the direct product of algebras S_i, each of which has a degree which is a power of some prime. If the original S is prime to M, it follows that each of the component algebras is also prime to M, by the theory of the local invariants.[4] Hence let S be an algebra whose index is a power of a prime l. Suppose that p is a prime divisor of F which is not ramified in K. Then, for every prime factor P of p in K, the Frobenius symbol $[K/F; P]$ has a unique representation $[K/F; P] = ([K'/F; P'], [K''/F; P''])$ where P', P'' denote the prime divisors induced by P in the subfields K', K'', respectively.[5] Therefore the degree $n(P) = n(p)$ of the local extension K_P/F_p is equal to the l. c. m. of the degrees $n(P')$, $n(P'')$ of $K'_{P'}/F_p$, $K''_{P''}/F_p$, respectively.

Now let $\rho(p)$ be an arbitrary invariant of an l-primary algebra S. Since $\rho(p)$ is a fraction with denominator $n(p)$, we may write

$$(1) \qquad \rho(p) = a(p)/n(p) \equiv a'(p)/n(P') + a''(p)/n(P'') \pmod 1$$

for p is, by assumption on S, prime to M. In general there will be several such representations (1) of $\rho(p)$ as the sum of fractions with denominators $n(P')$, $n(P'')$. We agree to select one such representation for each prime divisor p. By the fundamental relation[6] for the invariants of S we have

$$(2) \qquad \sum_p \rho(p) \equiv \sum_p [a'(p)/n(P')] + \sum_p [a''(p)/n(P'')] \equiv 0 \pmod 1.$$

---

[2] For the definition of these groups of algebras see S. MacLane and O. F. G. Schilling, "Normal algebraic number fields," *Transactions of the American Mathematical Society*, vol. 50 (1941), pp. 295-384, especially pp. 298-303.

[3] A. A. Albert, *Structure of Algebras*, American Mathematical Society Colloquium Publications, vol. 24 (1939), Theorem 18, p. 77.

[4] H. Hasse, "Die Struktur der R. Brauerschen Algebrenklassengruppe über einem algebraischen Zahlkörper," *Mathematische Annalen*, vol. 107 (1933), pp. 731-60. Quoted as HA. See in particular pp. 750, 752.

[5] H. Hasse, "Bericht über neuere Untersuchungen und Probleme der algebraischen Zahlkörper," *Jahresberichte der Deutsche Mathematische Vereinigung*, Ergänzungband 6 (1930), Part II, p. 7. Quoted as HB.

[6] HA. p. 750.

Since all invariants $\rho(p)$ have as denominators powers of l it follows from (2) that

$$(3) \qquad s' \equiv -\sum_p [a''(p)/n(P'')] \text{ and } s'' \equiv -\sum_p [a'(p)/n(P')]$$

have the same denominator l^m. Because there is only a finite number of prime divisors p giving non-zero contributions to s' and s'', it follows that l^m divides the l. c. m. of a finite number of local degrees $n(P')$, $n(P'')$. All prime divisors p under consideration are, by assumption, prime to M. Hence l^m is the order of elements τ' in Γ' and τ'' in Γ''. The elements τ', τ'' may be determined as follows. Pick a prime P'' involved in s', P' involved in s'' such that $n(P')$, $n(P'')$ are divisible by l^m. The associated Frobenius automorphisms σ'_0, σ''_0 are elements in the respective Galois groups. Then suitable powers of σ'_0, σ''_0 are elements τ', τ'' with the required properties.

We next make use of hypothesis (H). There exists at least one element $\sigma = (\sigma', \sigma'')$ in Γ whose components have orders divisible by l^m. The Tschebotareff density theorem [7] asserts that there are infinitely many prime divisors q in F which are prime to M such that

$$[K/F; Q] = \sigma = (\sigma', \sigma'') = ([K'/F; Q'], [K''/F; Q''])$$

where $Q \mid q$ in K and Q induces Q', Q'' in K', K'', respectively. By construction we have

$$l^m \mid n(Q') \qquad \text{and} \qquad l^m \mid n(Q'').$$

Consequently [8] s' and s'' can be realized as invariants of local algebras split by $K'_{Q'}/F_q$ and $K''_{Q''}/F_q$, respectively. Similarly, the fractions $a'(p)/n(P')$, $a''(p)/n(P'')$ are individually invariants of local algebras split by $K'_{P'}/F_p$ and $K''_{P''}/F_p$, respectively. By the definition (3) of s', s'',

$$(4) \qquad \begin{aligned} s' + \sum_p [a'(p)/n(P')] &\equiv 0 \pmod 1, \\ s'' + \sum_p [a''(p)/n(P'')] &\equiv 0 \pmod 1. \end{aligned}$$

Hence [9] there exist algebras S' and S'' split by K', K'', respectively, which have the invariants indicated in (4), where in particular s', s'' are the invariants at q. Relation (2) and definition (3) imply $s'' \equiv s' \pmod 1$. Consequently $S \sim S' \times S''$, for the associated invariants at q are equal, modulo 1.

In order to prove that condition (H) is necessary let us suppose that (H) is false. Then there exists a prime l, an integer m and elements τ' in Γ', τ'' in Γ'' with orders l^m such that for every element $\sigma = (\sigma', \sigma'')$ in Γ either

---

[7] HB. Part II, p. 133. [8] HA. p. 752. [9] HA. pp. 748, 752.

the order of σ' or that of σ'' is not divisible by l^m. From the elements τ', τ'' in the groups Γ', Γ'' we can then construct elements $\tau_1 = (\tau', \rho'')$ and $\tau_2 = (\rho', \tau'')$ in Γ. We choose ρ'' as any element in Γ'' which is mapped by $\mathbf{T}''$ into the element of Δ which is the image of τ' under $\mathbf{T}'$; similarly we choose ρ'. This construction can be done so that the elements τ_1, τ_2 both have order l^m, using the fact that hypothesis (H) is supposed to be false. By the Tschebotareff density theorem there exist prime divisors p_1 and p_2 of F whose corresponding Artin symbols [10] $(K/F; p_i) = \{\sigma[K/F; P_i]\sigma^{-1},\ P_i \mid p_i \text{ in } K, \sigma \text{ in } \Gamma\}$ have

$$\tau_1 \in (K/F; p_1) \qquad \text{and} \qquad \tau_2 \in (K/F; p_2).$$

This means that for any factor P_1 of p_1 in K the local degree $n(P_1)$ will be exactly the order l^m. We next construct an algebra S over F whose only invariants are

$$\rho(p_1; S) = 1/l^m \qquad \text{and} \qquad \rho(p_2; S) = -1/l^m.$$

Then S is split by K because the local degrees of K at p_1, p_2 are both equal to l^m, which is the local index of the algebra S.

Suppose now that the conclusion of the theorem held. Then $S \sim S' \times S''$, where S' and S'' are algebras split by K' and K'', respectively. We can assume here that the indices of S' and S'' are powers of the prime l, for otherwise we could write [11] $S' = S'_1 \times S'_2$ and $S'' = S''_1 \times S''_2$, where the indices of S'_1 and S''_1 are powers of l, and the indices of S'_2 and S''_2 are prime to l. One then would have the decomposition $S \sim S'_1 \times S''_1$, as asserted.

In the decomposition $S = S' \times S''$, the invariants must satisfy for every p the condition

$$\rho(p, S) \equiv \rho(p, S') + \rho(p, S'') \pmod 1.$$

From this it follows that $i(p, S)$ is the least common multiple of $i(p, S')$ and $i(p, S'')$, where $i(p, S)$ denotes the local index of S at p; that is, the smallest integer such that $i(p, S)\rho(p, S) \equiv 0 \pmod 1$. Since S' and S'' are split by K we have

$$i(p_1, S') \mid l^m \qquad \text{and} \qquad i(p_1, S'') \mid l^m.$$

Moreover, at least one of the indices must equal l^m because the original index $i(p_1, S) = l^m$. But now the Artin symbol $(K/F; p_1)$ is the class of conjugates of τ_1 in the group Γ, so that one may choose a prime factor P_1 of p_1 in K so that $\tau_1 = [K/F; P_1] = ([K'/F; P'_1],\ [K''/F; P''_1])$, where P'_1, P''_1 are

[10] HB. Part II, p. 6.

[11] By the decomposition theorem, A. A. Albert, *loc. cit.*

the prime multiples of P_1 in K', K'', respectively. We compare this equation with the construction of the element τ_1. We find $[K'/F; P'_1] = \tau'$ has order l^m and $[K''/F; P''_1] = \rho''$ has an order which is not divisible by l^m, for hypothesis (H) is assumed to be false. Therefore the local degree of K''/F at p_1 is not a multiple of l^m. Consequently $i(p_1; S'') = l^m$ is impossible and actually $i(p_1; S'') = l^{m_1}$ with $m_1 < m$. Hence we have

$$(5) \qquad \begin{aligned} i(p_1; S') &= l^m, \qquad i(p_1; S'') = l^{m_1}, \quad m_1 < m, \\ i(p_2; S') &= l^{m_2}, \qquad i(p_2; S'') = l^m, \quad m_2 < m. \end{aligned}$$

Since the sum of the local invariants of S' is $0 \pmod 1$ we have

$$(6) \qquad \rho(p_1; S') \equiv -\rho(p_2; S') - \sum_{q \neq p_1, p_2} \rho(q; S') \pmod 1.$$

Observing (5) we find that there exists a prime divisor q such that $i(q; S') = l^{m+\nu}$, $\nu \geqq 0$. But the invariant of S at q is equal to $0 \pmod 1$ by construction. Hence $0 \equiv \rho(q; S) \equiv \rho(q; S') + \rho(q; S'') \pmod 1$ implies $i(q; S'') = l^{m+\nu}$. Therefore the local degress of K', K'' at q are divisible by l^m. Next take a factor Q of q in K. Then $\sigma = [K/F; Q] = ([K'/F; Q'], [K''/F; Q''])$ where the orders of the components are divisible by l^m, Q' and Q'' the prime multiples of Q in K', K'', respectively. This contradicts the second part of the auxiliary assumption that hypothesis (H) is false.

We next discuss pairs of groups Γ', Γ'' satisfying hypothesis (H). Although we have not found a complete characterization of these groups we have analyzed various special cases. In particular, we show by examples that hypothesis (H) does not hold for all pairs of groups.

THEOREM 2. *Let Γ', Γ'' be two groups which have the same homomorphic image Δ. Suppose that $\{l\}$ is the set of all prime factors common to the orders of Γ', Γ''. Then hypothesis (H) holds for Γ', Γ'' if it holds for all pairs of l-Sylow groups Γ'_0, Γ''_0 of Γ', Γ'', respectively, which have the same image in Δ.*

Proof. Let τ', τ'' be elements of Γ', Γ'' which have the same order l^m. We have to prove the existence of an element (σ', σ'') in the subdirect product of Γ' and Γ'' whose components have orders divisible by l^m. Each of the elements τ', τ'' lies in some Sylow group Γ'_1, Γ''_0 of Γ', Γ'', respectively. The maps $\Gamma'_1\mathfrak{T}' = \Delta'_1$, $\Gamma''_0\mathfrak{T}'' = \Delta''_0$ of the Sylow groups into $\Delta = \Gamma'\mathfrak{T}' = \Gamma''\mathfrak{T}''$ are again [12] *l-Sylow* subgroups of Δ. The two groups Δ'_1, Δ''_0 are conjugate in Δ, thus $\Delta''_0 = \phi\Delta'_1\phi^{-1}$ for some element ϕ in Δ. Now pick an element ρ in Γ'

[12] See H. Zassenhaus, *Lehrbuch der Gruppentheorie*, Hamb. Math. Einzelschriften (1937), p. 100, Th. 2. Quoted as Z.

with $\rho T' = \phi$. Then $\Gamma'_0 = \rho\Gamma'_1\rho^{-1}$ is an l-Sylow group of Γ' the map of which covers all of Δ''_0. Thus, the groups Γ'_0, Γ''_0 have a common image Δ''_0 in Δ. The pair $(\rho\tau'\rho^{-1}, \tau'')$ has again components whose orders are divisible by l^m. We now use the assumption that hypothesis (H) holds for Γ'_0, Γ''_0. Then there exists an element (σ', σ'') of the subdirect product of Γ'_0, Γ''_0 which has the required properties. Since the latter group lies in the subdirect product of Γ' and Γ'' the hypothesis (H) holds.

We thus reduce the problem of deciding whether a pair of groups satisfies hypothesis (H) to groups of prime power order. For the latter class we prove

LEMMA 1. *If Γ', Γ'' are both regular*[13] *l-groups then hypothesis (H) holds.*

Proof. Suppose that l^{m+1} is the maximal order for the elements of a regular l-group Γ. Let Σ be the set of all elements not of maximal order in Γ. Then for every α in Σ we have $\alpha^{l^m} = 1$, while for the elements γ in the complementary set $\Gamma - \Sigma$

$$\gamma^{l^m} \neq 1 \qquad \text{and} \qquad \gamma^{l^{m+1}} = 1.$$

By a theorem of Hall[14] Σ is a subgroup of Γ. We return now to the groups Γ', Γ'' of the Lemma and consider the respective subgroups Σ', Σ''. By construction the elements of maximal order of Γ', Γ'' are found in the complementary sets $\Gamma' - \Sigma'$, $\Gamma'' - \Sigma''$. The maps $\Sigma'T'$ and $\Sigma''T''$ are subgroups of Δ. Suppose that both are proper subgroups. Then each subgroup $\Sigma'T'$, $\Sigma''T''$ contains at most half the number of elements in Δ. Since both homomorphic maps have the unit element of Δ in common, there must exist at least one element δ in Δ which lies neither in $\Sigma'T'$ nor in $\Sigma''T''$. There exist elements σ' in Γ' and σ'' in Γ'' which are mapped upon δ by the homomorphisms T', T''. These elements do not lie in Σ', Σ'' hence they have maximal order. The element (σ', σ'') of Γ has the required properties.

It remains to consider the case where $\Sigma'T' = \Delta$. Let σ', σ'' be elements in Γ', Γ'' which have maximal orders in the respective groups. Suppose that $\sigma'T' = \delta'$, $\sigma''T'' = \delta''$. By assumption on Σ', there exists then an element τ' in Σ' with $\tau'T' = (\delta')^{-1}\delta''$. By construction, the element $\sigma'\tau'$ has maximal order and $(\sigma'\tau')T' = \delta'(\delta')^{-1}\delta'' = \delta''$. Hence $(\sigma'\tau', \sigma'')$ is an element of Γ which has the desired properties. A similar argument may be applied in case $\Sigma''T'' = \Delta$.

---

[13] For the definition and properties of regular l-groups see P. Hall, "A contribution to the theory of groups of prime-power order," *Proceedings of the London Mathematical Society*, ser. 2, vol. 36 (1932), pp. 29-95, especially pp. 73-8.

[14] P. Hall, *loc. cit.* Theorem 4.26, p. 76.

Observe that we proved more than is required by the statement of the lemma. We actually proved the existence of an element σ whose components have maximal (possible) orders.

LEMMA 2. *The class of regular l-groups contains the following types of groups*:

(i) *every abelian group of prime power order,*

(ii) *every l-group, for odd prime l, whose commutator group lies in its center.*

Proof. We recall that an l-group of class c is regular whenever [15] $c < l$. Here the class of a group is the length of its lower central series. An abelian group has class 1; this gives case (i) above. In the case (ii) the class of the group is 2, and $2 < l$.

Suppose that Γ is represented as the subdirect product of two factor groups Γ' and Γ'' which satisfy the conditions of Lemma 1, 2 for their respective Sylow groups. We next observe that the direct product of regular groups is again a regular group. Then the Sylow subgroups of Γ satisfy the same conditions, for Γ is contained in the direct product of Γ' and Γ''. Conversely, suppose that the Sylow subgroups of Γ satisfy the above conditions. Let Γ' be any homomorphic image of Γ, say $\Gamma' = \Gamma/\Lambda$. Suppose the commutator groups of Γ, Γ' are Γ_0, Γ'_0, similarly let Z, Z' be the associated centers. Then $\Gamma'_0 = \Gamma_0 \bigcup \Lambda/\Lambda$ and $Z \bigcup \Lambda/\Lambda \leqq Z'$. Since $\Gamma_0 \leqq Z$ we have $\Gamma'_0 \leqq Z'$. Hence the validity of the hypotheses of Lemmas 1, 2 for Γ implies their validity for any two factor groups Γ', Γ'' such that Γ is the subdirect product [16] of Γ', Γ''.

We are now in a position to prove

THEOREM 3. *Let K/F be a normal field which is the join of t normal subfields K_i/F, $i = 1, 2, \cdots, t$, then $S = S_1 \bigcup \cdots \bigcup S_t$ where the S and S_i, $i = 1, \cdots, t$, are the groups of algebras prime to the module M of K/F and split by K, K_i, respectively, if either*

(i) *K is the direct join of the subfields K_i, or*

(ii) *all Sylow groups of the Galois group of K/F are regular.*

Proof. We shall prove the theorem by induction on the set of subfields K_i. For the first part of the theorem we observe that hypothesis (H) is finally

---

[15] P. Hall, *loc. cit.*, p. 4a, Definition 2.45 and p. 73, Corollary 4.13.

[16] Observe that the assumptions on the Sylow groups of Γ do not imply that Γ is a subdirect product of suitable factor groups. For an example consider the icosahedral group of $60 = 3 \cdot 4 \cdot 5$ elements. This group satisfies all assumptions on the Sylow groups but is not a subdirect product since it is a simple group.

satisfied. The assertion is true for $t = 1$. Suppose it is proved for any subset, say $K_1, \cdots, K_{t-1}$, that an algebra split by $K_1 \cup \cdots \cup K_{t-1}$ is the direct product of algebras split by $K_1, \cdots, K_{t-1}$. Then let $K_1 \cup \cdots \cup K_{t-1} = K'$ and $K_t = K''$. The Galois groups Γ', Γ'', are homomorphic maps of the group Γ. Thus their Sylow subgroups Γ'_0, Γ''_0 are maps of the Sylow subgroups of Γ. Hence both groups Γ'_0, Γ''_0 are regular groups.[17] Since Theorem 2 combined with Lemmas 1 and 2 and the remark at the beginning of this proof implies the validity of hypothesis (H) for the two cases under consideration, we may apply Theorem 1. Then the algebra S split by K is similar to the direct product $S' \times S''$ where $S' \times K' \sim K'$ and $S'' \times K'' \sim K''$. Since $S' \sim S_1 \times \cdots \times S_{t-1}$, $S'' = S_t$ we have $S \sim S_1 \times \cdots \times S_t$.

In the preceding arguments the assumption that the algebras concerned are prime to the modulus M cannot be dropped, as we now show by an example in which l is any given prime, F any algebraic number field, K a suitable extension of degree l^2 over F. First choose two distinct rational primes l_1 and l_2 in the arithmetic progression $1 + xl$, let p_i be a prime divisor of l_i in F, and let F_{p_i} be the corresponding complete (p-adic) field, for $i = 1, 2$. Each F_{p_i} has an unramified cyclic extension of degree l; on the other hand, the choice of the l_i insures that each F_{p_i} also has a ramified cyclic extension of the same degree (for example, the extension generated by the l-th root of any prime element). The Grunwald existence theorem[18] then gives two cyclic fields $K^{(i)}$ of prime degree l over F, such that p_i is ramified in $K^{(i)}/F$ and totally inert in $K^{(j)}/F$, for $i \neq j$. However, the join $K = K^{(1)} \cup K^{(2)}$ of these fields has local degree l^2 both at p_1 and p_2, so that there exists a division algebra split by K which has the invariants $1/l^2$, $-1/l^2$ at p_1, p_2, respectively.[19] This algebra is not similar to the direct product of algebras split by $K^{(1)}$ and $K^{(2)}$.

Furthermore, Theorem 3 is not true for arbitrary (non-normal) subfields K', K'', with $K' \cap K'' = F$, $K' \cup K'' = K$. This may be shown explicitly in the case when F is the rational field, K a field with the symmetric group of order 6, K' and K'' conjugate cubic subfields of K.

The following converse to the preceding theorems holds:

THEOREM 4. *If K/F is normal, and $K' \supset F$, $K'' \supset F$ are two subfields*

---

[17] For every homomorphic map of a regular group is regular. See P. Hall, *loc. cit.*, p. 74.

[18] W. Grunwald, "Ein algebraisches Existenztheorem für algebraische Zahlkörper," *Jour. für d. reine und ang. Math.*, vol. 169 (1933), pp. 103-7.

A. A. Albert, "On p-adic fields and rational division algebras," *Annals of Mathematics*, vol. 41 (1940), pp. 674-93; in particular pp. 688-90.

[19] HA. pp. 748, 752.

such that $S = S' \cup S''$, *for the associated group of algebra classes, then* $K = K' \cup K''$.

An indirect proof may be given, using the Tschebotaroff density theorem to construct suitable prime divisors inert in $K' \cup K''$ but not in K. We omit details.

We now wish to indicate the scope of the group theoretical arguments leading to Theorem 3, and in particular, to show that hypothesis (H) is not vacuous. Specifically, we exhibit for each prime l an irregular group for which hypothesis (H) holds and one for which it fails.

LEMMA 3. *Let* l *be an odd prime. Then the group* Γ' *with the generators* $Q_1, \cdots, Q_{l-1}$, R *and the relations* $Q_1{}^{l^2} = Q_2{}^l = \cdots = Q_{l-1}{}^l = R^l = 1$, $RQ_1R^{-1} = Q_1Q_2$, $RQ_2R^{-1} = Q_2Q_3, \cdots, RQ_{l-2}R^{-1} = Q_{l-2}Q_{l-1}$ *and* $RQ_{l-1}R^{-1} = Q_{l-1}Q_1{}^{-l}$ *is irregular.*

This group may be conveniently described as a group extension. Let A be the abelian group generated by the elements $Q_1, \cdots, Q_{l-1}$ of orders $l^2, l, \cdots, l$ and set $Q_l = Q_1{}^{-l}$, $Q_{l+1} = Q_2{}^{-1} = 1$. Then one may verify that the definition

$$Q_i{}^\alpha = Q_iQ_{i+1}, \qquad i = 1, 2, \cdots, l-1 \tag{1}$$

determines an automorphism α of A. By induction on k one then shows that

$$Q_1{}^{\alpha^k} = \prod_{i=1}^{k+1} Q_i{}^{d_i}, \quad \text{where } d_i = \binom{k}{i-1}, \text{ for } k = 1, \cdots, l. \tag{2}$$

In particular, since each binominal coefficient d_i with $k = l$ is divisible by l, one has $Q_1{}^{\alpha^l} = Q_1$. By (1) and induction on i, $Q_{i+1}{}^{\alpha^l} = Q_i{}^{\alpha^l\alpha}Q_i{}^{-\alpha^l} = Q_{i+1}$; hence α is an automorphism of order l. We may therefore regard Γ' as a group extension [20] of A by a cyclic group of order l, represented by an operator R with $R^l = 1$, $RQ_iR^{-1} = Q_i{}^\alpha$.

To determine the order of elements in Γ', let $NA = A^{1+\alpha+\cdots+\alpha^{l-1}}$ for any element of the group A. By (2) and by properties of binomial coefficients, one shows that $NQ_1 = 1$ and hence by (1) that $NQ_{i+1} = NQ_i{}^\alpha NQ_1{}^{-1} = 1$. Therefore $NA = 1$ for every A.

Each element of Γ' may be expressed as $B = R^jA$. If $j \not\equiv 0 \pmod l$, $B^l = R^{jl}NA = 1$, so B has order l. The elements of maximal order l^2 in Γ' are thus found in the subgroup A, and are exactly the elements given in the lemma. This implies that Γ' is irregular, for RQ_1 and R^{l-1} both have order l, whereas their product $R^{l-1}RQ_1 = Q_1$ has the larger order l^2. This cannot happen [21] in a regular group.

---

[20] Z., pp. 89-98.

[21] P. Hall, *loc. cit.*, p. 77, Th. 4.28.

From this group Γ' we construct a pair of groups for which hypothesis (H) fails. Let T' be the homomorphism $\Gamma' \to \Gamma'/A$ in which Γ' is mapped on a cyclic group Δ of order l, while all elements of maximal order in Γ' are mapped on the identity of Δ. Let $\Gamma''_1 = \{S\}$ be cyclic of order l^2, and let T''_1 be the homomorphism $\Gamma''_1 \to \Gamma''_1/\{S^l\}$. No element of maximal order l^2 in Γ''_1 is mapped on 1. This shows that hypothesis (H) fails in this case.

We may also use the irregular group Γ' to construct an example in which (H) holds; we let Γ''_2 be an abelian group generated by elements S_1, S_2 with respective orders l^2, l, while T''_2 is the homomorphism $\Gamma''_2 \to \Gamma''_2/\{S_1\} = \Delta$. The maps of elements of maximal orders in Γ', Γ''_2 then have an element of Δ in common; namely, the identity of Δ.

Since l is an odd prime we can realize [22] the above groups Γ', Γ''_1, as Galois groups of fields K'_1, K''_1 over any algebraic number field F. By Theorem 1 we thus have fields K'_1, K''_1 with $K_1 = K'_1 \bigcup K''_1$, $S > S' \bigcup S''$.

In the case $l = 2$, exactly similar examples may be constructed. Instead of the group Γ' of Lemma 3 we use the generalized quaternion group (Z., p. 110) of order 2^n, $n \geqq 4$. This group is generated by two elements α, β with $\alpha^{2^{n-1}} = 1$, $\beta^2 = \alpha^{2^{n-2}}$, $\beta\alpha\beta^{-1} = \alpha^{-1}$. The elements of maximal order in this group are the elements α^j with $j \not\equiv 0 \pmod 2$. This group is irregular because $\alpha^2\beta$ and $\beta^{-1}\alpha$ both have order 4, while their product α^3 has order $2^{n-1} > 4$.

HARVARD UNIVERSITY,
UNIVERSITY OF CHICAGO.

---

[22] See for example H. Reichardt, "Konstruktion von Zahlkörpern mit gegebener Galoisgruppe von Primzahlpotenzordnung," *Jour. für d. reine und ang. Math.*, vol. 177 (1937), pp. 1-5.

T. Tannaka, "Ueber die Konstruktion der galoisschen Körper mit vorgegebener p-Gruppe," *Tohoku Math. Journ.*, vol. 43 (1937), pp. 252-260.

NORMAL EXTENSIONS OF RELATIVELY COMPLETE FIELDS.*

By O. F. G. SCHILLING.

This paper is a discussion of some relations between the Hilbert theory and the existence problem for normal extensions of relatively complete fields. The Hilbert theory describes the subfields of a given normal field by means of the arithmetic inherent in the fields in question. The results of this theory may be interpreted as necessary conditions for the existence of normal extensions over a base field. We propose the question whether such conditions are sufficient for the solution of existence problems. To answer this query it is necessary to construct explicitly normal extensions taking advantage of information obtained by the Hilbert theory. A. A. Albert, O. Ore and the author [1] have investigated independently such problems for the classical fields of p-adic numbers. In this article we take general relatively complete fields as reference fields. A field is termed relatively complete if Hensel's Lemma holds for the prime ideal of non-units with respect to the given valuation.[2] The normal extensions involved are mostly of infinite degree over the ground field. This generality leads to an understanding of the mechanism underlying the structure of normal extensions. We have purposely selected relatively complete fields rather than complete fields as base fields since relative completeness can always be established for infinite extensions.[3]

In the first part of the paper we develop the Hilbert theory for arbitrary normal extensions L over a relatively complete field F with value group Γ and *perfect* residue class fields $\mathfrak{F}$. We prove that L contains a unique maximal unramified field L_I whose residue class field coincides with that of L. The field L contains, relative to L_I, a radical extension L_R, the ramification field, which is determined by the structure of the value group $\Gamma(L)$ of L. As a matter of fact, the value group $\Gamma(L_R)$ of L_R consists of all those elements of $\Gamma(L)$ whose orders modulo Γ are prime to the characteristic χ of F. Moreover, the field L_I contains all roots of unity whose orders occur as orders of the elements in the factor group $\Gamma(L_R)/\Gamma$. In case L contains only elements whose degrees over F are prime to χ we find $L = L_R$. Thus normal extensions

* Presented to the Society, Dec. 29-31, 1941. Received October 29, 1941.

[1] See [1], [14], [20]. The numbers in square brackets refer to the bibliography at at the end of the paper.

[2] The definition of a relatively complete field was first given by Ostrowski in [16].

[3] For simple examples see [12], [15].

L of this type determine uniquely an extension $\mathfrak{L}$ of the residue class field $\mathfrak{F}$ and an extension $\Gamma(L) = \Delta$ of the original value group Γ.

The second part of the paper is devoted to the examination of the above results for the construction of normal fields with preassigned algebraic and arithmetic properties. We start with a normal extension L of F and a group $\Delta \supset \Gamma$. It is assumed that the degrees of the elements in L/F and the orders of the elements in Δ/Γ are prime to χ. One of our main results states that a normal field L with residue class field $\mathfrak{L}$ and value group Δ exists if, and only if, $\mathfrak{L}$ contains all characters belonging to the finite subgroups of Δ/Γ. The proof of this theorem furnishes the basis for the description of all finite groups G which can be realized as Galois groups of normal fields L/F. We present complete results for relatively complete fields F whose residue class fields $\mathfrak{F}$ are finite. In this special instance we succeed in determining an abstract discrete group which acts as a universal covering group for all realizable Galois groups G. The universal group is described solely in terms of the field F and the value group Γ of F.

Let $F = \{a, b, c, \cdots\}$ be a field on which a valuation [4] ϕ with the value group $\Gamma = \{\alpha, \beta, \gamma, \cdots\}$ is defined. Suppose that O, P are the ring of integers and the prime ideal of non-units, respectively, defined in F with respect to ϕ. We shall assume once and for all that the residue class field $\mathfrak{F} = O/P$ is a perfect field. The characteristic χ of $\mathfrak{F}$ shall be termed the characteristic of ϕ.

DEFINITION [5] 1. *The field F is said to be relatively complete with respect to the valuation ϕ if each congruence*

$$f(x) \equiv g_0(x)h_0(x) \pmod{P}$$

with

$$(g_0(x), h_0(x)) \equiv 1 \pmod{P}$$

implies

$$f(x) = g(x)h(x)$$

with

$$g(x) \equiv g_0(x) \pmod{P}, \quad h(x) \equiv h_0(x) \pmod{P}$$

where $f(x)$, $g_0(x)$, $h_0(x)$, $g(x)$, $h(x)$ lie in $O[x]$ and the degree of $g(x)$ is not larger than the degree of $g_0(x)$.

---

[4] For the basic facts of the general valuation theory see, for example, [8], [19].

[5] This definition differs slightly from the one given in [16], p. 296.

In the sequel we shall suppose that all fields F considered are relatively complete [6] with respect to some valuation ϕ.

An immediate consequence of Definition 1 is the following

LEMMA [7] 1. *If u is a unit of F whose residue class $\mathfrak{u}$ is the m-th power of an element $\mathfrak{v}$ in $\mathfrak{F}$ where m is relatively prime to the characteristic χ of ϕ, then u is the m-th power of a unit v in F.*

The method used for the construction of the quotient field of an integral domain may be employed [8] to prove

THEOREM 1. *For every abelian ordered group Γ and Steinitz number N there exists a unique (to within isomorphisms) smallest group $N^{-1}\Gamma \supseteq \Gamma$ such that each equation $n\xi = \gamma$ for $n \mid N$, γ in Γ, has a unique solution ξ in $N^{-1}\Gamma$. The group $N^{-1}\Gamma$ has the same rank [9] as Γ.*

We shall say that the elements $\gamma_1, \cdots, \gamma_n$ of Γ are a set of representatives for the factor group $\Gamma/p^\lambda\Gamma$, p a rational prime, if each γ in Γ has a unique representation $\gamma = \sum_{i=1}^{n} a_i\gamma_i + p^\lambda\gamma'$ where the a_i are positive representatives for the residues mod p^λ and γ' lies in Γ.

Using this definition and the homomorphism principle of the group theory one can readily prove

LEMMA 2. *If $[\Gamma : p\Gamma] = p^h$ then $[\Gamma : p^\lambda\Gamma] = p^{h\lambda}$ and each set of representatives for $\Gamma/p\Gamma$ is a set of representatives for $\Gamma/p^\lambda\Gamma$.*

Now let K be a finite algebraic extension of degree n over F. Suppose that $N_{K/F}A = NA$ denotes the norm of an element A in K taken from K to F. Then the field K is relatively complete with respect to the valuation [10] ϕ_K defined by $\phi_K(A) = (1/n)\phi(NA)$. The finiteness of the degree n implies that (i) the residue class degree $f = [\mathfrak{K} : \mathfrak{F}]$ is a divisor of n, and (ii) the index $[\Gamma(K) : \Gamma]$ is finite and $\Gamma(K) \subseteq (1/n)\Gamma$.

It is noteworthy that also every infinite algebraic extension L/F is

---

[6] There exist fields which are relatively complete without being complete. The examples of [12] and [15] can easily be generalized to valuations of arbitrary rank.

[7] For the proof of this lemma see [19], pp. 561-2.

[8] For a different type of proof see [11], Ch. I, 1. This proof has to be supplemented in our case so as to yield the invariance of the rank.

[9] For the definition of the rank see [8]. We emphasize that the rank of an ordered abelian group is equal to the order type of the set of isolated subgroups. Thus all subgroups of the additive group of real numbers have the same (order) rank.

[10] The value group of ϕ_K will be denoted by $\Gamma(K)$, similarly $\mathfrak{K}$ stands for the residue class field of K.

relatively complete with respect to a unique extension Φ of ϕ. For a proof of this remark it suffices to observe that the coefficients of the polynomials $f(x)$, $g_0(x)$, $h_0(x)$, involved in Definition 1, determine a finite extension K/F L/F. Thus, the existence of the polynomials $g(x)$, $h(x)$ can be established by referring to the analogous assertion for finite algebraic extensions. To determine the value of an element in L it suffices to find its value for any finite subfield M/F of L. Thus the factor group $\Gamma(L)/\Gamma$ contains only elements of finite order. Finally, the residue class field $\mathfrak{L}/\mathfrak{J}$ contains the residue class fields $\mathfrak{K}$ for all finite extensions K/F contained in L/F. These results may, in particular, be applied to any one of the isomorphic algebraic completions $\bar{F}$ of F. Consequently, we may assume [11] that all value groups and fields considered are contained in the respective sets of a fixed algebraic completion $\bar{F}$ of F.

An infinite separable algebraic extension L/F will be termed normal if every finite subset of elements in L lies in a finite (separable) normal extension K/F contained in L/F. By a simple well-ordering argument we can suppose [12] that L is given as the join $\bigcup_\nu K_\nu$ of finite normal extensions K_ν/F where the set $\{K_\nu\}$ is a lattice with respect to joins and intersections. Each approximating set $\{K_\nu\}$ of L/F determines then an isomorphic representation of the Galois group $G(L/F)$ by means of vectors $\{g_\nu\}$ whose components g_ν are elements in the Galois groups $G_\nu = G(K_\nu/F)$. A vector $\{g_\nu\}$ represents an element in $G(L/F)$ if and only if g_ν is induced by g_μ whenever $K_\nu \subset K_\mu$. It is not difficult to establish the Galois correspondence between the subfields M/F of L/F and the subgroups of $G(L/F)$ which are closed with respect to any one of the topologies of $G(L/F)$ defined by the approximating sets $\{K_\nu\}$.

DEFINITION 2. *A subgroup S of $G(L/F)$ is called an everywhere dense subgroup if the closure of S with respect to the topology of $G(L/F)$ is equal to $G(L/F)$.*

THEOREM 2. *The finite subfields of L/F can be described by the subgroups of finite index in any everywhere dense subgroup S of $G(L/F)$.*

For a proof [13] one has to observe that the elements of S define on every finite subfield K/F the same automorphisms as the elements of $G(L/F)$. The

---

[11] Since the property of an element being integral is transitive and since the prime ideal P of F is divisible by exactly one prime ideal of F, we agree to use the symbol P ambiguously for the prime ideals of the various algebraic extensions $K \supset F$.

[12] For the Galois theory of infinite extensions see [3], [6] and the supplement in [22].

[13] See [20].

arguments used for denumerable extensions can immediately be generalized to non-denumerable fields.

DEFINITION 3. *A finite or infinite normal extension L/F is termed a p-extension if every element of L/F satisfies an equation whose degree is a power of a fixed rational prime p.*

Let L^*/F be the join of all cyclic extensions Z/F of degree p in a given p-extension L/F. Then L^*/F belongs to the closure of the subgroup generated by the commutator group $G(L/F)'$ of $G(L/F)$ and the p-th power of elements in $G(L/F)$. For a proof we observe first that $G(L^*/F)$ is an abelian group which is the limit of cyclic groups of order p. Hence $G(L/L^*)$ contains $G(L/F)'$. Since each element h of $G(L^*/F)$ is the limit of elements whose p-th powers are equal to unity it follows that the p-th power of any element g in $G(L/F)$ mapped upon h lies in $G(L/L^*)$. Hence $G(L/L^*)$ contains the closure of $\{G(L/F)', g^p, g \text{ in } G(L/F)\}$. Conversely, let L^{**} be the subfield of L/F belonging to the closure of $\{G(L/F)', g^p, g \text{ in } G(L/F)\}$. Then L^{**} is an abelian extension which is the join of cyclic extensions of degree p. Since L^* is maximal we have $L^* \supseteq L^{**}$. Hence $G(L/L^*) \subseteq G(L/L^{**})$.

THEOREM 3. *Let L/F be a normal p-extension. If $[L^*:F] = p^h$ then $G(L/F)$ can be generated by h elements $g_1, \cdots, g_h$; that is to say, every element of $G(L/F)$ is the limit of finite products in $g_1, \cdots, g_h$. The number h is the smallest number of generators.*

Proof. Suppose that $\{K_\nu\}$ is an approximating set of L/F such that $K_\nu \supseteq L^*$. Such a set exists since L^* has finite degree over F. If G_ν denotes the Galois group of K_ν/F then $\lim_\nu G_\nu = G(L/F)$. By the theory of finite groups [14] it follows that each group G_ν can be generated by h elements. Let $\tilde{G}_\nu$ be the totality of all sets [15] of h generators of G_ν. These sets are finite for each group G_ν is finite. If G_ν is mapped homomorphically upon G_μ, i.e. $L_\mu \subset L_\nu$, then $\tilde{G}_\nu$ is mapped into $\tilde{G}_\mu$. Hence $\lim_\nu \tilde{G}_\nu$ is a non-void subset of $G(L/F)$. Take, then, h elements $g_1, \cdots, g_h$ in the limit set. These elements may be selected so as to determine a generating set in each $\tilde{G}_\nu$. Hence the infinite discrete group generated by $g_1, \cdots, g_h$ in $G(L/F)$ is an everywhere dense subgroup. Consequently, by Theorem 4, the elements of $G(L/F)$ are limits of finite products of $g_1, \cdots, g_h$. Similarly it follows that at least h generators are needed.

---

[14] See [23].

[15] See [21].

DEFINITION 4. *A separable field L/F is called an unramified extension if $[K:F] = [\mathfrak{K}:\mathfrak{F}]$ for all finite subfields K/F in L.*

LEMMA 3. *The set of unramified extensions over F is a lattice with respect to joins and intersections of fields as operations.*

Proof.[16] Since each infinite unramified extension can be approximated by finite unramified extensions it suffices to prove the lemma for the latter. We observe that the residue class field $\mathfrak{F}_\infty$ of the separable algebraic completion F_∞ of F is an algebraic completion of $\mathfrak{F}$. Then the arguments of Ostrowski in [16], Section 44, pp. 340-341 furnish a proof for the theorem.[17] We remark that we only need the relative completeness of the base field F to generalize Ostrowski's theory of inertial quantities.

LEMMA 4. *If L/F is an unramified extension then $\Gamma(L) = \Gamma$ for the associated value groups.*

Proof. Since each γ in $\Gamma(L)$ is the value of an element A in L it suffices to prove the assertion for the fields $K = F(A)$. Then, by Definition 4, $[\mathfrak{K}:\mathfrak{F}] = [K:F] = n$. Hence, as a consequence of the relative completeness, $K = F(B)$ where B maps upon a primitive element of $\mathfrak{K}/\mathfrak{F}$. Then $A = \sum_{i=0}^{n-1} a_i B^i$, a_i in F. On multiplying A by a suitable integer a of F we may suppose that the $a_i a$ are integers where $\phi(a_j a) = 0$ for at least one subscript j. Then $\phi(Aa) \equiv \phi(A) \pmod{\Gamma}$. We find $\phi(Aa) = 0$ for otherwise $Aa \equiv 0 \equiv \sum_{i=0}^{n-1} (a_i a)B^i \pmod{P}$, contrary to the assumption for B.

THEOREM 4. *The field L/F contains a unique maximal unramified extension W/F such that $\mathfrak{W} = \mathfrak{L}$. There is a 1-1 correspondence between the unramified subfields of L/F and the subfields of $\mathfrak{L}/\mathfrak{F}$; in particular, the normality of L/F implies the normality of W/F and $G(W/F) = G(\mathfrak{L}/\mathfrak{F})$.*

Proof. By the customary device of approximating an infinite extension by finite extensions it suffices to establish the assertions of the lemma for finite fields K/F. Let $[\mathfrak{K}:\mathfrak{F}] = f$ and suppose that the primitive element $\mathfrak{A}$ of $\mathfrak{K}/\mathfrak{F}$ is a root of the equation $x^f + \mathfrak{a}_1 x^{f-1} + \cdots + \mathfrak{a}_f = 0$. Pick any set of elements a_j in F with $a_j \bmod P = \mathfrak{a}_j$. Then $f(x) = x^f + a_1 x^{f-1} + \cdots a_f = 0$ is irreducible in F and has, by the relative completeness of K, a linear factor

[16] For these and subsequent arguments consult [5], [7], [12], [13], [16].

[17] We omit the detailed study of the theory of inertial quantities as developed first in [16] for valuations of rank one.

$x - A$ in K with $A \bmod P = \mathfrak{a}$. Hence $\{F(A)\} \bmod P = \mathfrak{F}(\mathfrak{a})$; thus $F(A)$ is an unramified extension of F. Next let M be another unramified subfield of K/F, $M \subset F(A)$. Then comparison of degrees shows $[\{M \cup F(A)\} \bmod P : \{F(A)\} \bmod P] = 1$, contrary to the assumption for M. The remaining assertions of the lemma are direct consequences of the uniqueness of the field $F(A) = W$.

As special cases we note

COROLLARY 1. *If $\{K_\nu\}$ is an approximating set of L/F then $\mathfrak{L} = \bigcup_\nu \mathfrak{K}_\nu$.*

COROLLARY 2. *If $\mathfrak{F}$ is a finite field with q elements then there exists for each positive integer n exactly one unramified extension W_n of degree n over F. The field W_n is the cyclotomic extension defined by the $(q^n - 1)$-th roots of unity.*

LEMMA 5. *Let L/F be an algebraic extension such that each element of L is the root of an equation of degree prime to χ, then $L = F$ if $\Gamma(L) = \Gamma$, $\mathfrak{L} = \mathfrak{F}$.*

Proof. It suffices to consider finite extensions for $L = \{K_\nu\}$, $\Gamma(L) = \Gamma$, $\mathfrak{L} = \mathfrak{F}$ imply $\Gamma(K_\nu) = \Gamma$, $\mathfrak{K}_\nu = \mathfrak{F}$. The hypotheses of the lemma imply that a finite extension K/F has a primitive element A with $A^n + a_1 A^{n-1} + \cdots + a_n = 0$ where $n = [K:F]$ and $\phi(a_n) = 0$. Then $nt(A) \equiv 0 \pmod P$, t the trace for K/F. Consequently $A \equiv 0 \pmod P$, for $A \equiv a \pmod P$ with a in F, contrary to the assumption on A. Hence $n = 1$.

DEFINITION 5. *An algebraic separable extension L/F is termed totally ramified if $\mathfrak{L} = \mathfrak{F}$.*

THEOREM 5. *If L/F is a totally ramified extension such that each element of L is the root of an equation of degree prime to χ then L is a radical extension of F.*

Proof. As usual it suffices to prove the assertion for finite totally ramified extensions K/F. Suppose the theorem is proved for fields K/F for which $\Gamma(K)/\Gamma$ is a cyclic group of order m. In other words $K = F(A)$ where $A^m = a$ lies in F and $\{\phi(A), \Gamma\} = \Gamma(K)$. Let, for a general finite extension K/F, $\gamma_1, \cdots, \gamma_r$ be a complete set of representatives for an independent system of cosets of $\Gamma(K)/\Gamma$. Next pick A_j, $j = 1, \cdots, r$, in K such that $\phi(A_j) = \gamma_j$ and $A_j^{m_j} = a_j$ in F where m_j is the order of $\gamma j \bmod \Gamma$. Such a selection is always possible by a proper normalization of the A_j, observing Lemma 1, by units of F. Then $K = F(A_1) \cup \cdots \cup F(A_r)$ as follows by Lemma 5. Finally, suppose that $\Gamma(K)/\Gamma = \{\delta, \Gamma\}/\Gamma$ is a cyclic group of

order m. As before we select an element A of K with $\phi(A) = \delta$ and $A^m = a$ in F. Then $F(A) = K$ for $\{F(A)\} \bmod P = K \bmod P$ and $\Gamma(F(A)) = \Gamma(K)$. We remark that the element a is determined by the value groups $\Gamma(K)$ to within m-th powers of elements in F. In particular, any unit u of F may be used as a multiplier so that some solution of $x^m = au$ still generates K.

THEOREM 6. *Let L/F be a totally ramified extension. Suppose that $\Delta \supset \Gamma$ is a subgroup of $\Gamma(L)$ such that the orders of all elements in Δ/Γ are prime to χ. Then there exists*[18] *a unique (smallest) $L(\Delta)$ of L with $\Gamma(L(\Delta)) = \Delta$.*

Proof. By the customary method of approximation, observing that $\Delta = \bigcup_\nu \Delta_\nu$ where the groups Δ_ν/Γ are finite, we reduce the proof to the finite case. The proof of the preceding theorem implies the existence of a field $L(\Delta) = F(A_1, \cdots, A_r)$, $\{\phi(A_1), \cdots, \phi(A_r), \Gamma\} = \Delta$, with $\Gamma(L(\Delta)) = \Delta$. Suppose that $\bar{L}$ is another subfield of L with $\Gamma(\bar{L}) = \Delta$. Then $\bar{L}$ contains a field $L = F(\bar{A}_1, \cdots, \bar{A}_r)$ where $\bar{A}_j^{m_j} = A_j^{m_j} = a_j$, by the concluding remark in the proof of Theorem 5. Hence $L(\Delta)$ and $\bar{L}$ are conjugate fields in L/F. Consequently $L(\Delta) \bigcup \bar{L}$ is obtained from $L(\Delta)$ by the adjunction of roots of unity. Therefore $\mathfrak{L} \supseteq \{L(\Delta) \bigcup \bar{L}\} \bmod P \supset \mathfrak{J}$, contrary to the assumption on L.

We state, without proof,

COROLLARY 3. *Let $\Gamma(L)'$ be the (unique) subgroup of $\Gamma(L)$ consisting of all elements in $\Gamma(L)$ whose orders* mod Γ *are prime to χ. Then the totally ramified extension L/F contains a unique minimal subfield L'/F such that $\Gamma(L') = \Gamma(L)'$. The degree (over L') of each element of L which does not lie in L', is a power of χ.*

COROLLARY 4. *If the totally ramified extension L/F is normal then the fields $L(\Delta)$, determined by Theorem 6, are abelian. Moreover, all totally ramified extensions $\{L_\mu/F\}$ so that each cyclic subgroup of $\Delta(L_{\mu_1})/\Gamma$ is isomorphic to some cyclic subgroup of $\Delta(L_{\mu_2})/\Gamma$, $\mu_1 \neq \mu_2$ in $\{\mu\}$, are abelian if and only if one of the fields in the set $\{L_\mu/F\}$ is normal.*

Proof. By the arguments employed for the proof of Theorems 5 and 6, it suffices to show that a cyclic factor group Δ/Γ for $\Delta \subset \Gamma(L')$ gives rise to a cyclic extension. The fact that L/F is totally ramified combined with the

[18] We must envisage the possibility $\Gamma(L') = \Gamma(L)$ though $L' \subset L$. For an example see [18], pp. 682-3.

normality implies that all roots of an equation $x^m = a$, $\{(1/m)\phi(a), \Gamma\} = \Delta$ and a in F, lie in L. Hence F must contain the m-th roots of unity. Consequently $L(\Delta) = F(a^{1/m})$ is cyclic. The second statement of the corollary follows now readily.

In order to determine the explicit structure of the Galois group of a normal extension L/F it is useful to develop a generalization of the Hilbert theory.[19]

DEFINITION 6. *Let g be an element of the Galois group G of L/F. An element A of L is called a semi-invariant with respect to g if* $A^{1-g} = U(A, g) \equiv 1 \pmod{P}$.

DEFINITION 7. *The totality of all automorphisms s in G for which all units of L/F are semi-invariants, is called the inertial set I of L/F.*

LEMMA 6. *The inertial set I of L/F is an invariant closed subgroup of G.*

Proof. The first part of the assertion follows by explicit computation of the effect of the elements $s_1 s_2$, s_i^{-1}, $g s_i g^{-1}$ with s_1, s_2, in I, g in G, on the units of L. For the second part we observe that $I_\nu \bigcup G(K_\nu/K_\mu)/G(K_\nu/K_\mu) \cong I_\mu$ for any two fields $K_\nu \subset K_\mu$ of an approximating set $\{K_\nu\}$ of L. Consequently, every vector $\{g_\nu\}$ of G with g_ν in I_ν is an element of I. Conversely, each s in I induces in each field K_ν an element s_ν of I_ν. Hence $I = \lim_\nu I_\nu$, for the topology of G, so that I is closed.

THEOREM 7. *The factor group G/I is isomorphic to the Galois group of the residue class field $\mathfrak{L}$ over $\mathfrak{F}$.*

Proof. We first define the elements g of G as operators on L/F. Let $\mathfrak{A}^g = (U^g \bmod P)$ where U is some unit of L representing the residue class $\mathfrak{A}$. The definition implies that $\mathfrak{A}^g$ does not depend on the particular unit U. For $V \bmod P = U \bmod P$ implies $V = UW$ where $W \equiv 1 \pmod{P}$. Thus each element g defines a unique automorphism of L/F since the elements of F are not changed by g. In particular we have $\mathfrak{A}^s = \mathfrak{A}$ for elements s in I, by definition of the inertial group. Consequently G/I is isomorphic to a subgroup of $G(\mathfrak{L}/\mathfrak{F})$. Now let $\mathfrak{A}$ be an arbitrary element of $\mathfrak{L}$. We want to show that $\mathfrak{F}(\mathfrak{A})$ has its normal closure in $\mathfrak{L}$. Let $\mathfrak{A}^n + \mathfrak{a}_1\mathfrak{A}^{n-1} + \cdots + \mathfrak{a}_n = 0$ be the defining equation of $\mathfrak{A}$ with $\mathfrak{a}_j$ in $\mathfrak{F}$. Select these elements a_j in F

[19] For the subsequent definitions and methods consult [4], [5], [7], [8], [12], [13], in particular [8]. A set of more restrictive definitions may be set up by demoting some of our lemmas and theorems to definitions.

with $a_j \bmod P = \mathfrak{a}_j$. Since $a_n \neq 0$ we have $\phi(a_n) = 0$. The equation $f(x) = x^n + a_1 x^{n-1} + \cdots + a_n = 0$ is irreducible in F and has one root A in L with $A \bmod P = \mathfrak{A}$, for both F and L are relatively complete fields. Since L/F is normal the polynomial $f(x)$ factors completely in L. All conjugates of A are units in L for $\phi(A) = (1/n)\phi(a_n) = 0$. Consequently all conjugates of $\mathfrak{F}(\mathfrak{A})$ lie in $\mathfrak{L}$, that is $\mathfrak{L}$ is normal over $\mathfrak{F}$. Moreover, all the conjugates of $\mathfrak{A}$ in $\mathfrak{L}/\mathfrak{F}$ are maps of the conjugates of A. Hence all automorphisms of $G(\mathfrak{L}/\mathfrak{F})$ are induced by elements of G. Whence $G(\mathfrak{L}/\mathfrak{F}) = GI$.

DEFINITION 8. *The field L_I/F belonging to the inertial group I of L/F is called the inertial field of L/F.*

THEOREM 8. *The inertial field L_I/F of L/F is unramified; it is the join of the inertial fields of any approximating set for L/F. The value group $\Gamma(L_I)$ is equal to Γ. Moreover L_I/F is maximal, that is, it contains all unramified subfields of L/F.*

Proof. Let $\{K_\nu\}$ be an approximating set of L/F. Then, by Corollary 1, $L = \bigcup_\nu K_\nu$. By Theorem 4 there exist uniquely determined fields W_ν/F in K_ν/F with $[W_\nu : F] = [\mathfrak{W}_\nu : \mathfrak{F}] = [\mathfrak{K}_\nu : \mathfrak{F}]$ such that $W = \bigcup_\nu W_\nu$ is normal over F and $G(W/F) \cong G(\mathfrak{L}/\mathfrak{F})$. The elements s in I induce the identical automorphism of W/F, for if some element of W would be moved by an element s then s would not induce the identical automorphism of $\mathfrak{L}/\mathfrak{F}$. Consequently $W \subseteq L_I$. Finally $W = L_I$ for $G(W/F) \cong G/I$. The construction of W implies the maximality of L_I and $W = L_I = \bigcup_\nu (K_\nu)_I$ for any approximating set $\{K_\nu\}$ of L. The assertion $\Gamma(L_I) = \Gamma$ is a special case of Lemma 4.

DEFINITION 9. *The set of all automorphisms in the Galois group G of L/F for which all elements of L are semi-invariants is called the ramification set R of L/F.*

LEMMA 7. *The inertial set R of L/F is an invariant closed subgroup of G.*

Proof. The proof may be carried out exactly as was the proof of Lemma 6. We have to replace I_ν by R_ν and I by R. Observe that each r in R induces an automorphism r_ν in the set R_ν for an approximating field K_ν of L.

DEFINITION 10. *The field L_R/F belonging to the ramification group R of L/F is called the ramification field of L/F.*

THEOREM 9. *The ramification field L_R/F is an abelian extension of the*

inertial field L_I/F. The orders of the elements in $\Gamma(L_R)/\Gamma$ are prime to χ and the degree of each element of L_R over L_I is prime to χ.

Proof. Let A be an element of L which is a semi-invariant for an automorphism s in I. Then all elements AU, where U denotes an arbitrary unit of L, are also semi-invariants with respect to s for $(AU)^s \equiv A^s U^s \equiv AU \pmod{P}$, by definition of the inertial group. Suppose now that $\bar{a}$ is an arbitrary element of $\Gamma(L_R)/\Gamma$. Let A be an element of L_R such that $\phi(A) \bmod \Gamma = \bar{a}$. Then each element s of I determines by $A^{s-1} \equiv c(\bar{a}, s) \pmod{P}$ a unique element $c(\bar{a}, s)$ in the multiplicative group $\mathfrak{L}^*$ of the residue class field $\mathfrak{L}$. The element $c(\bar{a}, s)$ is different from 0 since $\phi(A) = \phi(A^s)$. By the remark made at the beginning of this discussion it follows that each product AU, U a unit, gives rise to the same residue class $c(\bar{a}, s)$. Moreover, on multiplying A by an element $b \neq 0$ in L_I, we obtain $c(\bar{a}, s)$ as the residue class for Ab since $b^s \equiv b \pmod{P}$. Hence $c(\bar{a}, s)$ depends solely on the residue class $\bar{a}$ and the element s.

Let $G(\bar{a})$ be the closed set of elements in I which have all elements A with $\phi(A) \bmod \Gamma = \bar{a}$ as semi-invariants. This set $G(\bar{a})$ is an invariant subgroup of I. We now associate to every s in I the element $c(\bar{a}, s)$, for fixed $\bar{a}$. Then each $\bar{a}$ determines by $s \to c(\bar{a}, s)$ a character of I, as follows by explicit computation. By definition of $G(\bar{a})$ we find that $c(\bar{a}, s)$ is exactly the identity in $G(\bar{a})$. Thus $c(\bar{a}, s)$ is a character on $I/G(\bar{a})$, and $I/G(\bar{a})$ is isomorphic to the group which is generated by the $c(\bar{a}, s)$ for variable s in I. Repeated application of s to A implies $A^{s^k} \equiv A c(\bar{a}, s)^k \pmod{P}$. There exists a smallest positive integer $m(\bar{a}, s) = m$ for which $c(\bar{a}, s)^m = 1$. To find an upper bound M for m we argue as follows. Let M be the order of $\bar{a}$ in $\Gamma(L_R)/\Gamma$. Then $1 = c(\bar{a}M, s) = c(\bar{a}, s)^M$. Whence $m \mid M$. Now let P_χ be the prime field of $\mathfrak{F}$. Then the characters $c(\bar{a}, s)$, s arbitrary in I, generate a cyclotomic finite extension $\mathrm{P}_\chi < \bar{a} >$ of P_χ. The group $X(\bar{a}, I)$ generated by the elements $c(\bar{a}, s)$, s variable in I, is a subgroup of the multiplicative group of elements of finite order in $\mathrm{P}_\chi < \bar{a} >$. We distinguish the two cases, $\chi = \infty$ and $\chi < \infty$. In the first case it follows by Dirichlet's theorem on units, that $X(\bar{a}, I)$ is cyclic since it is a subgroup of the group of all roots of unity in $\mathrm{P}_\chi < \bar{a} >$. In the second case $X(\bar{a}, I)$ is a cyclic group for the multiplicative group $\mathrm{P}_\chi < \bar{a} >^*$ is cyclic of order $\chi^N - 1$ where $N = [\mathrm{P}_\chi < \bar{a} > : \mathrm{P}]$. Therefore the group $I/G(\bar{a})$ is in either case a finite cyclic group whose order is prime to χ.

The ramification group R is by definition the largest subgroup of I for which all elements of L are semi-invariants. Since $R \subseteq G(\bar{a})$ we have $R \subseteq \bigcap_{\bar{a}} G(\bar{a})$. Now let s in $\bigcap_{\bar{a}} G(\bar{a})$. This element s has $c(\bar{a}, s) = 1$ for all $\bar{a}$.

Consequently all elements are semi-invariants with respect to s. Whence $R = \bigcap_{\bar{a}} G(\bar{a})$.

Now let $L < \bar{a} >$ be the subfield of L/F which belongs to $G(\bar{a})$ and let $n(\bar{a}) = [I : G(\bar{a})] = [L < \bar{a} > : L_I]$. Consider the join $\bigcup_{\bar{a}} L < \bar{a} >$ for all $\bar{a}$. This field is equal to L_R by the Galois theory for $R = \bigcap_{\bar{a}} G(\bar{a})$. Thus we find that the ramification field L_R/L_I is the join of the finite fields $L < \bar{a} > /L_I$ whose degrees $n(\bar{a})$ are prime to χ. Hence $\Gamma(L_R) = \bigcup_{\bar{a}} \Gamma(L < \bar{a} >)$. By Theorem 6 and the proof of Lemma 7 it follows that L_R belongs to the subgroup $\Delta = \Gamma(L_R) \supseteq \Gamma_I = \Gamma$; moreover, the Galois group I/R of L_R/L_I is abelian. Finally we remark that $G(L < \bar{a} > /L_I) = I/G(a)$ and $\Gamma(L < \bar{a} >)/\Gamma = \{\Gamma, \bar{a}\}/\Gamma \cong I/G(\bar{a})$.

THEOREM 10. *Let $\Gamma(L)'$ be the subgroup of all elements in $\Gamma(L)$ whose orders* mod Γ *are prime to χ. Then L_R is the smallest subfield of L for which $\Gamma(L_R) = \Gamma(L)'$. The fields M with $L_I \subseteq M \subseteq L_R$ are in* 1-1 *correspondence with the subgroups Δ of $\Gamma(L)'$ which contain Γ.*

Proof. By Theorem 6 the field $L' \supseteq L_I$ is uniquely determined by $\Gamma(L)'$. Now let $\bar{a}$ be an arbitrary element of $\Gamma(L)'/\Gamma$. As before we construct the group $G(\bar{a}) \subseteq I$. By definition of R we have $G(\bar{a}) \supseteq R$. Since $\bigcap_{\bar{a}} G(\bar{a})$, taken as in the proof of Theorem 9, is equal to R we have *a fortiori* $\bigcap_{\bar{a}} G(\bar{a}) = R$ where the intersection is taken for all $\bar{a}$ in $\Gamma(L'/\Gamma$. Hence $L' = L_R$ by the Galois theory. Next we observe that $\Gamma(M) \subseteq \Gamma(L)'$ for a field M with $L_I \subseteq M \subseteq L_R$. Conversely each group Δ with $\Gamma(L_R) \supseteq \Delta \supseteq \Gamma$ determines uniquely a field $M(\Delta)$ with $L_R \supseteq M(\Delta) \supseteq L_I$ as follows from Theorem 9.

THEOREM 11. *If I/R is a locally compact group satisfying the second axiom of denumerability then the character group of I/R is isomorphic to the factor group $\Gamma(L_R)/\Gamma = \Gamma(L)'/\Gamma$.*

Proof. We use a different interpretation of the function $c(\bar{a}, s)$. We now keep s fixed and vary $\bar{a}$. Then $c(\bar{a} + \bar{\beta}, s) = c(\bar{a}, s)c(\bar{\beta}, s)$ for $\bar{a}$, $\bar{\beta}$ in $\Gamma(L)'/\Gamma$. Hence $c(\bar{a}, s)$ is a character of $\Gamma(L)'/\Gamma$. Consequently I/R is isomorphic to a subgroup of the character group of $\Gamma(L)'/\Gamma$. The properties of R imply that for every s in I, not in R, there is at least one $\bar{a}$ with $c(\bar{a}, s) \neq 1$. Consequently only the identity of I/R furnishes the identity on $\Gamma(L)'/\Gamma$. Now let $\{M_\nu\}$ be an approximating set of L_R/L_I. Then

$\lim_{\nu} I/G(L_R/M_\nu) = I/R$. Next we have $G(M_\nu/L_I) = I/G(L_R/M_\nu)$. For if $\Gamma(M_\nu) = \{\alpha_1^{(\nu)}, \cdots, \alpha_{t_\nu}^{(\nu)}, \Gamma\}$ and $\alpha_j^{(\nu)} \bmod \Gamma = \bar{a}_j^{(\nu)}$ we have $G(M_\nu/L_I) \cong I/\bigcap_j G(\bar{a}_j^{(\nu)}) \cong \Gamma(M_\nu)/\Gamma$. As before we may interpret $I/G(L_R/M_\nu)$ as the character group of $\Gamma(M_\nu)/\Gamma$. Now $\bigcup_\nu \Gamma(M_\nu) = \Gamma(L_R)$, whence $\lim_\nu \Gamma(M_\nu)/\Gamma = \Gamma(L_R)/\Gamma$. Hence it follows by the theory of characters of abelian groups,[20] that the character group of $\Gamma(L_R)/\Gamma$ is isomorphic to I/R.

THEOREM 12. *If the ramification field L_R/F contains a cyclic extension Z/L_I then L contains all roots of unity belonging to the character group of $G(Z/L_I)$.*

Proof. If Z/L_I is a finite extension of degree n then L_I contains, by Corollary 4, the n-th roots of unity. For the infinite case, observe that Z/L_I may be approximated over L_I by a denumerable set $\{Z_j\}$. Let $[Z_j : Z_{j-1}] = n_j$. Then L_I contains the roots of unity of order $\prod_{\nu=1}^{j} n_\nu$. The latter constitute the characters of $G(Z/L_I)$ which in turn is isomorphic to the closure of the additive group of the natural integers with respect to the principal ideals $(\prod_{\nu=1}^{j} n_\nu)$ as neighborhoods of zero. Recalling the properties of relatively complete fields it finally follows that the above roots of unity lie in L; thus they lie in $\mathfrak{L}$.

COROLLARY 5. *If L/F is an infinite normal extension such that the group of all roots of unity in L is finite and L_R is an infinite extension of L_I, then Γ has infinite rational rank*[21] *over the group of integers.*

Proof. Let N be the order of the group of roots of unity in L_I. Suppose that α is an element of $\Gamma(L_R)$ with exact order $n(\alpha)$ relative to Γ. Then the $n(\alpha)$-th roots of unity lie in L_I, so that $n(\alpha) \mid N$. By previous considerations it follows that there is exactly one field $L_I(A)$ in L_R with $\phi(A_a) = \alpha$ and $A_a^{n(a)} = a_a$ in L_I. The element a_a is uniquely determined by $\alpha \bmod \Gamma$ to within $n(\alpha)$-th powers of elements of L_I. Since all $n(\alpha)$ are divisors of N the field L_R/L_I must contain infinitely many independent cyclic extensions of prime degree p dividing N. Since Γ contains no elements of finite order the

---

[20] See [17].

[21] The rational rank of an (ordered) abelian group Γ over the additive group of integers is defined as the number of linearly independent elements of Γ over the integers. In general, the order rank is less than the rational rank. See [9] for a special case.

factor group $\Gamma/p\Gamma$ must be infinite. Hence there are infinitely many elements in Γ which are independent over the additive group of integers.

The structure of the normal field L over its ramification field L_R is now easy to determine. Theorem 5 applied to L relative to L_R implies that each element of L which is not in L_R satisfies over L_R an equation whose degree is a power of χ. Thus we have

THEOREM 13. *Each element of L is the root of an equation over L_R whose degree is a power of χ, provided $\chi \neq \infty$. If $\chi = \infty$ then $L = L_R$, in other words $R = 1$.*

We state without proof [22] that a more detailed investigation of the structure of L/L_R for $x \neq \infty$ shows that $G(L/L_R)$ is a (solvable) χ-group. If L is an infinite extension over L_R then $G(L/L_R)$ is a zero-dimensional topological group whose structure closely resembles the structure of the classical Lie groups.

The Hilbert theory, as we developed it, may be viewed as follows. The existence of a normal field L/F whose Galois group is supposed to be known *in abstracto* implies certain special properties and relations between the Galois group and the arithmetical structure of the field L/F and its subfields . Thus, the Hilbert theory may be interpreted as a set of necessary conditions for the existence of normal extensions with prescribed Galois groups. In the subsequent investigations we shall discuss sufficient conditions for the existence problem. Theorems 1, 4, 7 and 8 describe completely the connection between the normal extensions of the residue class field $\mathfrak{F}$ and the class of inertial fields of the normal extensions of F.

In the sequel we shall suppose once and for all that the elements of extensions L/F, K/F satisfy equations over F whose degrees are prime to the characteristic χ of the residue class field $\mathfrak{F}$.

THEOREM 14. *Suppose that $\Gamma \neq p\Gamma$ for the value group of the base field F, p a rational prime. If F does not contain the p-th roots of unity then every cyclic extension of F is unramified, and conversely.*

Proof.[23] Suppose that F does not contain a primitive p-th root of unity ζ. Let Z be a cyclic extension of degree p^m over F. If Z were a ramified extension then $[Z : Z_I] \geqq p$ by Theorem 5 and the proof of Theorem 11, and the assump-

[22] We omit in this paper a detailed study of the theory of higher ramification groups and their group-theoretical structure. This theory has no significance for the solution of the special existence problem.

[23] A different proof, without using the Hilbert theory, may be set up using the analysis of [1], pp. 209-11.

tion on Γ. Then Z_I must contain the primitive $[Z:Z_I]$-th roots of unity. Consequently, the degree $[Z_I:F]$ has to be divisible by a factor of $p-1$ for $[F(\zeta):F] \mid p-1$. This is impossible. Conversely, suppose that F has only cyclic unramified extensions. We find that F has ramified cyclic extensions of degree p if ζ lies in F. For take a in F with $\phi(a)$ not in $p\Gamma$, then $F(a^{1/p})$ is a ramified field, contrary to the hypothesis.

LEMMA 8. *If $\Gamma = p\Gamma$ then F has at most unramified normal p-extensions.*

Proof. By Theorem 13 we have $L = L_R$. If L were a proper extension of L_I then $\Gamma(L) \supseteq \Gamma$. This is impossible for $\Gamma(L) \subseteq \bigcup_\nu p^{-\nu}\Gamma = \Gamma$. Hence F admits only unramified p-extensions.

Arguments similar to the ones above imply

LEMMA 9. *A field F has cyclic completely ramified extensions of degree n if and only if the n-th roots of unity lie in F and $\Gamma/n\Gamma$ contains at least one element of order n.*

Let $F^{(p)}/F$ be the join of all normal p-extensions. We shall prove that in certain cases the Galois group of any finite normal p-extension has a finite number of generators which is bounded by arithmetic invariants of the base field F..

THEOREM 15. *Suppose that F contains the p-th roots of unity. The Galois group of $F^{(p)}/F$ has a finite number of generators if and only if the factor groups $\mathfrak{J}^*/\mathfrak{J}^{*p}$ and $\Gamma/p\Gamma$ are finite.*

Proof. We first remark that $[F^*:F^{*p}] = [\Gamma:p\Gamma][\mathfrak{J}^*:\mathfrak{J}^{*p}]$ provided the first index is finite. Consider on F^* and F^{*p} the homomorphism $a \to \phi(a)$. Then $[F^*:F^{*p}] = [\Gamma:p\Gamma][U:U^p]$ where U is the group of units in F. We remark that the subgroup of F^{*p} which is mapped upon 1 is exactly equal to U^p for the fact that U^p is in F^{*p} with $\phi(b) = 0$ implies b in U with $b = c^p$, c in U. Next apply to $[U:U^p]$ the homomorphism $U \to U \bmod P = \mathfrak{U}$. Then $[U:U^p] = [\mathfrak{J}^*:\mathfrak{J}^{*p}]$ since, by Lemma 1, every unit U which is $\equiv 1 \pmod P$ is a p-th power.

Now suppose that $G(F^{(p)}/F)$ has a finite number of generators, say h. Then, the join $\bigcup_\nu Z_\nu$ of all cyclic subfields $Z \subset F^{(p)}$ of degree p over F is a finite extension of F for $G(F^{(p)}/F)$ has a finite number of generators. Since $[\bigcup_\nu Z_\nu : F] = [F^*:F^{*p}]$ we see that both $[\Gamma:p\Gamma]$ and $[\mathfrak{J}^*:\mathfrak{J}^{*p}]$ must be finite.

Conversely, suppose that $[\Gamma:p\Gamma]$ and $[\mathfrak{J}^*:\mathfrak{J}^{*p}]$ are finite. Let

$L = \bigcup_p Z_v$. Let L_I be the inertial field of L. Then $[L_I : F] = [\{L_I\} \bmod P : \mathfrak{F}] = [\mathfrak{F}^* : \mathfrak{F}^{*p}]$ by Theorems 4, 7, 8. Since $\Gamma(L_I) = \Gamma$ and $\Gamma \subseteq \Gamma(L_I) \subset p^{-1}\Gamma = \Gamma(L)$ it follows from the proof of Theorem 11 that $[L : L_I] = [\Gamma(L) : \Gamma]$ is finite. Thus $[L : F]$ is finite.

THEOREM 16. *If F does not contain the p-th roots of unity then $G(F^{(p)}/F) = G(\mathfrak{F}^{(p)}/\mathfrak{F})$; in particular $G(F^{(p)}/F)$ has a finite number of generators if and only if $\mathfrak{F}$ has a finite number of independent cyclic extensions of degree p.*

Proof. We consider again the field $L = \bigcup_v Z_v$ in $F^{(p)}$. Then Theorem 14 and Lemma 8 imply that none of the fields Z_v/F is ramified. Hence L is an unramified extension and thus $\mathfrak{L} = \bigcup_v \mathfrak{Z}_v$, by Lemma 3 and Theorem 4. Next we observe that $F_I^{(p)}$ is a normal p-extension; as such it cannot contain the primitive p-th roots of unity. Consequently $F^{(p)} = F_I^{(p)}$ for the existence of a cyclic field of degree p over $F_I^{(p)}$ would imply, by Lemma 9, that all p-th roots of unity lie in $F_I^{(p)}$. Hence $F^{(p)}$ is an unramified extension and consequently $G(F^{(p)}/F) = G(\mathfrak{F}^{(p)}/\mathfrak{F})$ by Theorems 4 and 8.

Suppose now that $G(F^{(p)}/F)$ has a finite number of generators. Then $\bigcup_v Z_v$ is a finite extension of F and $\bigcup_v Z_v$ is the join of all cyclic extensions of degree p over $\mathfrak{F}$ for each cyclic extension $\mathfrak{Z}/\mathfrak{F}$ of degree p gives rise to a uniquely determined cyclic extension Z in $F^{(p)}$, as follows by Theorem 4. Hence $[\bigcup_v Z_v : F] = [\bigcup_v \mathfrak{Z}_v : \mathfrak{F}] = p^h$ and $G(F^{(p)}/F) = G(\mathfrak{F}^{(p)}/\mathfrak{F})$ can be generated by h elements. Conversely Theorems 3 and 4 imply that $G(F^{(p)}/F)$ has a finite number of generators if $\mathfrak{F}$ admits only a finite number of independent cyclic extensions of degree p.

LEMMA 10. *If F contains the p-th roots of unity and $G(F^{(p)}/F) = G(\mathfrak{F}^{(p)}/\mathfrak{F})$ then $\Gamma = p\Gamma$.*

Proof. The assumption implies $F^{(p)} = F_I^{(p)}$. Consequently $\Gamma(F^{(p)}) = \Gamma$ by the proof of Theorem 11. Therefore $\Gamma = p\Gamma$ for $\Gamma \supset p\Gamma$ would imply $F^{(p)} \supset F_I^{(p)}$.

As a corollary to the preceding results we next prove

THEOREM 17. *Suppose that the residue class field $\mathfrak{F}$ of F is a finite field. Then the following statements hold for the number of generators of the Galois group $G(F^{(p)}/F)$.*

(A) *If F does not contain the p-th roots of unity then $G(F^{(p)}/F) = G(\mathfrak{F}^{(p)}/\mathfrak{F})$ is isomorphic to the additive group of p-adic integers.*

(B) *If* $\Gamma = p\Gamma$ *and* F *contains the* p^m*-th roots of unity,* $1 \leqq m < k$, k *some positive integer, then* $G(F^{(p)}/F) = G(\mathfrak{F}^{(p)}/\mathfrak{F})$ *is isomorphic to the additive group of p-adic integers.*

(C) *If* $[\Gamma : p\Gamma] = p^h$ *and* $\mathfrak{F}$ *contains all* p^m*-th roots of unity,* $m = 1, 2, \cdots$, *then* $G(F^{(p)}/F)$ *is isomorphic to the direct sum of* h *groups which are isomorphic to the additive group of p-adic integers.*

(D) *If* $\mathfrak{F}$ *contains the* p^m*-th roots of unity,* $1 \leqq m < k$, k *some positive integer, and* $[\Gamma : p\Gamma] = p^h$, *then* $G(F^{(p)}/F)$ *can be generated by* $h + 1$ *elements.*

Proof. We first determine the structure of $G(\mathfrak{F}^{(p)}/\mathfrak{F})$. From the theory of Galois fields it follows: (i) $G(\mathfrak{F}^{(p)}/\mathfrak{F})$ is an infinite cyclic group if $\mathfrak{F}$ contains a finite group of p^m-th roots of unity. If the p^m-th roots of unity lie in $\mathfrak{F}$ but not the p^{m+1}-th roots of unity, then $\mathfrak{F}$ has for each power p^n exactly one cyclic extension of degree p^n. It follows then that $G(\mathfrak{F}^{(p)}/\mathfrak{F})$ is isomorphic to the additive group of p-adic integers; (ii) $G(\mathfrak{F}^{(p)}/\mathfrak{F})$ is equal to unity if $\mathfrak{F}$ contains all p^m-th roots of unity, $m = 1, 2, \cdots$.

Statement (A) follows by Theorem 16, (B) by Lemma, and (D) from Theorem 15. To prove (C) we observe that F has no unramified extension of degree p. Using Theorem 7 we prove that $G(F^{(p)}/F)$ has the required structure. Pick h representatives $\alpha_1, \cdots, \alpha_h$ for the generators of $\Gamma/p\Gamma$. Then select h elements $a_1, \cdots, a_h$ in F with $\phi(a_j) = \alpha_j$. The radicals a_j^{1/p^m}, $m = 1, 2, \cdots$, $j = 1, \cdots, h$, generate the field $F^{(p)}$ over F since the elements $p^{-m}\alpha_j$ generate $\bigcup_m p^{-m}\Gamma \supseteq \Gamma$ for Γ has no elements of finite order. By construction $F(\{a_j^{1/p^m}\})$, $m = 1, 2, \cdots$, is an infinite cyclic extension whose Galois group is isomorphic with the additive group of p-adic integers. Hence $G(F^{(p)}/F)$ has the required structure. We observe that $F^{(p)}/F$ does not depend on the selection of the elements a_j for all p^m-th roots of unity in F lie in F, by assumption on F and Lemma 1.

Our final goal is the complete description of $G(F/^{(p)}F)$ for case C. Before doing so we shall prove some general results concerning the extensibility of an imbedding of a given normal field K/F in larger normal fields M/F. From now on we shall suppose that the ground field F contains a primitive p-th root of unity ζ.

THEOREM 18. *Let* Z/F *be a cyclic extension of degree* p^m *with the residue class field* $\mathfrak{W}/\mathfrak{F}$, *let* $f = [W : F]$ *and* W *the inertial field* Z_I *of* Z/F *with* $[Z : W] = p^m f^{-1} = e$. *Then* Z/F *can be imbedded in a cyclic field* Z'/F *of degree* p^{m+1} *if and only if there exists an element* $\mathfrak{c}$ *in* $\mathfrak{W}$, *such that* $\zeta \bmod P(\mathfrak{N}\mathfrak{c})^e$ *where* $\mathfrak{N}$ *denotes the norm from* $\mathfrak{W}$ *to* $\mathfrak{F}$.

Proof. In order that a cyclic field Z'/F with the required properties exists it is necessary and sufficient that $\zeta = N_{Z/F}(A)$ with A in Z. Since $\phi(\zeta) \neq 0$ the element A has to be a unit. We may then write $A = cU$ where c lies in W and $A \equiv c \pmod P$. Thus U is a unit in Z, depending on the choice of c, with $U \equiv 1 \pmod P$. Then $N_{Z/F}A = N_{Z/F}cN_{Z/F}U = (N_{W/F}c)^e N_{Z/F}U$. Passing to the residue class field $\mathfrak{W}$ of Z we find $\zeta \equiv (N_{W/F}c)^e \pmod P$ or $\zeta \bmod P = (\mathfrak{N}\mathfrak{c})^e$ where $c \bmod P = \mathfrak{c}$. Conversely the existence of an element $\mathfrak{c}$ with the above property implies in turn that ζ is the norm of an element in Z. We select an element B in W such that $B \bmod P = \mathfrak{c}$. Then $\zeta(N_{W/F}B)^{-e} = \zeta N_{Z/F}B^{-1} \equiv u \equiv 1 \pmod P$. By Corollary 1 there exists a unit $v \equiv 1 \pmod P$, lying in F such that $v^{ef} = u$. Then $\zeta N_{Z/F}B^{-1} = v^{ef}$ and hence $\zeta = N_{Z/F}Bv$.

COROLLARY 6. *Suppose that $\mathfrak{F}$ is a Galois field with q elements. Then a cyclic field Z/F of degree p^m with ramification index e is imbeddable in a cyclic field Z' of degree p^{m+1} over F if and only if $e \mid (q-1)p^{-1}$.*

Proof. Let Ω be a fixed primitive (q^f-1)-st root of unity in $\mathfrak{W}/\mathfrak{F}$. Then $\mathfrak{N}\Omega = \omega$ is a primitive $(q-1)$-st root of unity of $\mathfrak{F}$. Let $\zeta \bmod P = \zeta_\phi$ be given as ω^r with $r = (q-1)p^{-1}$. Then field Z'/F exists, by Theorem 17, if and only if $\zeta_\phi = \omega^r = (\mathfrak{N}\Omega^x)^e = \omega^{xe}$, where $\mathfrak{c} = \Omega^x$. Hence the necessary and sufficient condition means that the congruence $xe \equiv r \pmod{q-1}$ must have a solution x. By elementary number theory this is the case if and only if $(e, q-1) \mid r$. Let μ be chosen so that $p^\mu \mid q-1$, $p^{\mu+1} \nmid q-1$ and let $e = p^\epsilon$. Then the condition may be written in the form $(p^\epsilon, p^\mu) \mid r$ or $p^\delta \mid p^{\mu-1}$ where $\delta = \text{Min}(\epsilon, \mu)$. Hence $\epsilon \leqq \mu - 1$ or $e \mid (q-1)^{-1}p$.

Let W be a normal unramified extension of degree f over F and let $G(W/F) = \{S_1, \cdots, S_f\}$. Let Δ be an extension of the value group Γ such that $[\Delta : \Gamma]$ is finite and prime to χ. Suppose that $\delta_1, \cdots, \delta_r$ is a linearly independent basis of Δ/Γ such that $n_i\delta_i = \alpha_i$ lie in Γ, $i = 1, \cdots, r$, n_i the minimal orders of the $\delta_i \bmod \Gamma$. Pick then s elements a_i in F with $\phi(a_i) = \alpha_i$.

THEOREM 19. *An extension*[24] $K = W((a_1C_1)^{1/n_1}, \cdots, (a_rC_r)^{1/n_r})$ *with units C_i in W is normal of degree $[\Delta : \Gamma]f$ over F and has W as inertial field and Δ as value group if and only if* (i) *the n_i-th roots of unity, $i = 1, \cdots, r$,*

[24] For a special case see [2]. After the completion of this paper MacLane and the author discovered that analogous results can be developed for function fields. Some of the results obtained in [11] can be proved for relatively complete fields. We quote especially the theorems on the explicit structure of the possible Galois groups as group extensions. The latter possibility is indicated by Theorem 19.

lie in $\mathfrak{W}$ *and* (ii) $(C_i^{S_{\nu-1}})$ mod P *are* n_i*-th powers in* $\mathfrak{W}$, $i = 1, \cdots, r$, $\nu = 1, \cdots, f$.

Proof. Let K/F be a normal extension with W as inertial field and $\Gamma(K) = \Delta$. Then $[K:F] = f[\Delta:\Gamma]$ by the general ramification theory. Let $A_1, \cdots, A_r$ be a set of elements in K, with $\phi(A_i) = \delta_i$. Then the elements $A_i^{n_i} a_i^{-1}$ are units of K. Suppose $A_i^{n_i} a_i^{-1} \equiv C_i \pmod{P}$ with C_i in W, then $A_i^{n_i} a_i^{-1} C_i^{-1} = U_i \equiv 1 \pmod{P}$. By Corollary 1 we find a unit V_i in K with $V_i^{n_i} = U_i$. Let $B_i = A_i V_i^{-1}$. By definition of B_i we have $B_i^{n_i} = a_i C_i$ and $K = W(B_1, \cdots, B_r)$. Next we observe that the fields $W(B_i)$ are normal over F. Since K is normal over F the equations, $x^{n_i} = (a_i C_i)^{S_\nu}$, $\nu = 1, \cdots, r$, factor completely in K. We have $(a_i C_i)^{S_\nu} = a_i C_i \cdot C_i^{S_{\nu-1}}$. Hence all the quantities $(C_i^{S_{\nu-1}})^{1/n_i}$ lie in K and generate over W inertial extensions. Consequently, by assumption on K, the elements $(C_i^{S_{\nu-1}})$ mod P lies in $\mathfrak{W}^{*n_i}$ or $(C_i \bmod P)^{S_{\nu-1}}$ in $\mathfrak{W}^{*n_i}$. Moreover, by Theorem 11, the n_i-th roots of unity lie in $\mathfrak{W}$.

Conversely let a_i be given and suppose that the equations $y^{n_i} = (C \bmod P)^{S_{\nu-1}}$, $i = 1, \cdots, r$, $\nu = 1, \cdots, f$, are soluble in $\mathfrak{W}$ for given $\mathfrak{C}_i$ of $\mathfrak{W}$. By (i) all roots must lie in $\mathfrak{W}$. Then the equations $y^{n_i} = C_i^{S_{\nu-1}}$ factor completely in W, by Lemma 1. Consider the fields $W([a_i C_i)^{S_\nu}]^{1/n_i}$ for variable automorphisms S_ν in $G(W/F)$. By assumption (ii) on the unit C_i, all these fields coincide. Thus $K_i = W([a_i C_i]^{1/n_i})$ is a normal extension of F. We have, by construction, $[K_i:F] = n_i f$; W is the inertial field of K_i and $\Gamma(K_i) = \{\delta_i, \Gamma\}$. The join $K_1 \bigcup \cdots \bigcup K_r$ has the required properties.

We next determine how, for a given extension K/F, the units C_i depend on the elements a_i in F. Let $a_i = a_i u_i$ be any set of elements in F with $\phi(a'_i) = \alpha_i$. Then $K_i = W(B'_i)$ where $(B'_i)^{n_i} = a'_i C'_i$. Since B_i and B'_i are generating radicals we have, by the Kummer theory, $a'_i C'_i = a_i C_i X_i^{n_i}$ where $\phi(X_i) = 0$, X_i in W. Hence $C'_i = C_i u_i^{-1} X_i^{-n_i}$. Conversely $(a'_i C'_i)^{1/n_i}$ generate the field K_i if $C'_i = C_i u_i^{-1} X_i^{-n_i}$ where $a'_i = a_i u_i$.

Now let $a_1, \cdots, a_r$ be fixed once and for all and let $C_1, \cdots, C_r$ be a set of units in W such that $C_i^{S_{\nu-1}} = D_i (S_\nu)^{n_i}$, $i = 1, \cdots, r$, $\nu = 1, \cdots, f$, in W. As we have already seen the field $K = W(\{(a_i C_i)^{1/n_i}\})$ has Δ as value group and W as inertial field. On multiplying C_i by elements $v_i Y_i^{n_i}$ where the v_i are units in F and Y_i arbitrary elements of W^* we find $(C_i v_i Y_i^{n_i})^{S_{\nu-1}} = D_i (S_\nu)^{n_i} (Y_i^{S_{\nu-1}})^{n_i}$. Hence $K' = W(\{(a_i C_i v_i Y_i^{n_i})^{1/n_i}\}) = W(\{(a_i C_i v_i)^{1/n_i}\})$ is a normal extension of F which has again Δ as value group and W as inertial field. By the Kummer theory it follows that $K = K'$ if and only if v_i lies in W^{*n_i}. Consequently a field basis $\{a_i\}$ and a fixed set of units C_i with $C_i^{S_{\nu-1}}$ in W^{*n_i} gives rise to as many distinct fields K/F as there are distinct residue

classes of the unit group of F modulo the n_i-th powers of units in W. Since units which are $\equiv 1 \pmod{P}$ are n_i-th powers it follows that the distinct extensions K/F are in one-to-one correspondence with the number of residue classes of $\mathfrak{F}^*$ modulo $\mathfrak{W}^{*n_i}$ in $\mathfrak{W}^*/\mathfrak{W}^{*n_i}$.

In order to determine the Galois group $G(K/F)$ it suffices to determine the Galois groups $G_i = G(K_i/F)$ of the normal subfields K_i/F of K/F. For, if the groups $G_i = \{g_i\}$ are known then $G(K/F)$ consists of all the r-tuples $[g_1, \cdots, g_r]$ of elements g_i in G_i such that g_i, $i = 1, \cdots, r$, induce the same automorphism on the common subfield W/F. To simplify the notation we shall drop the subscript i and consider a typical field $W([aCu]^{1/n})$ where $C^{S_\nu - 1} = D(S_\nu)^n$ for all S_ν in $G(W/F)$. Let $\mathbf{T}$ be a generating automorphism of $G(K/W)$. Then $G(K/F)/\{\mathbf{T}\} = G(W/F)$, i.e., $G(K/F)$ is a group extension of $\{\mathbf{T}\}$ by $G(W/F)$. Each automorphism S_ν of $G(W/F)$ gives rise to exactly n distinct extensions to $G(K/F)$. Let $\mathfrak{S}_\nu$, $\nu = 1, \cdots, [W:F]$, be representatives of the cosets S_ν. In order to determine one extension $\mathfrak{S}_\nu$ we may proceed as follows. Letting $A = (aCu)^{1/n}$ we have $A^{\mathfrak{S}_\nu} = AU_\nu$ where U_ν is a unit of K for $\phi(A^{\mathfrak{S}_\nu}) = \phi(A)$. Since $\mathfrak{S}_\nu$ is an automorphism we have $(A^n)^{\mathfrak{S}_\nu} = (A^{\mathfrak{S}_\nu})^n$. Explicit computation shows $(A^{\mathfrak{S}_\nu})^n = A^n U_\nu{}^n = aCuU_\nu{}^n$ and $(A^n)^{\mathfrak{S}_\nu} = (aCu)^{S_\nu} = aC^{S_\nu}u$. Whence $U_\nu{}^n = C^{S_\nu - 1}$. Now $C^{S_\nu - 1} = D(S_\nu)^n$. Hence $U_\nu = D(S_\nu)\zeta^j$, $0 \leqq j < n - 1$, where ζ is a primitive n-th root of unity. Let $D(S_\nu)$ be a fixed solution of $y^n = C^{S_\nu - 1}$ for each S_ν. We shall select $\mathfrak{S}_\nu$ so that $A^{\mathfrak{S}_\nu} = AD(S_\nu)$. Next we agree to use a fixed n-th root of unity ζ. Then $A^{\mathbf{T}} = A\zeta^{a(\mathbf{T})}$ and $\zeta^{S_\nu} = \zeta^{b(S_\nu)}$. In order to determine the permutation rules for the group G it suffices to compute these rules for the elements $\mathfrak{S}_\nu$ and $\mathbf{T}$. We find $A^{\mathfrak{S}_\nu \mathbf{T}} = (A^{\mathbf{T}})^{\mathfrak{S}_\nu} = (A\zeta^{a(\mathbf{T})})^{\mathfrak{S}_\nu} = AD(S_\nu)\zeta^{a(\mathbf{T})\,b(S_\nu)}$ and $A^{\mathbf{T}\mathfrak{S}_\nu} = (A^{\mathfrak{S}_\nu})^{\mathbf{T}} = (AD(S_\nu))^{\mathbf{T}} = AD(S_\nu)\zeta^{a(\mathbf{T})}$. Consequently $\mathfrak{S}_\nu \mathbf{T} = \mathbf{T}^{b(S_\nu)}\mathfrak{S}_\nu$. Next let $S_\nu S_\mu S_\nu{}^{-1} S_\mu{}^{-1} = S_\rho$. Then the extensions $\mathfrak{S}_\nu \mathfrak{S}_\mu \mathfrak{S}_\nu{}^{-1} \mathfrak{S}_\mu{}^{-1}$ and $\mathfrak{S}_\rho$ must differ by a power of $\mathbf{T}$ since both lie in the same coset mod $\{\mathbf{T}\}$. We find $A^{\mathfrak{S}_\nu \mathfrak{S}_\mu \mathfrak{S}_\nu{}^{-1} \mathfrak{S}_\mu{}^{-1}} = AD(S_\nu S_\mu S_\nu{}^{-1} S_\mu{}^{-1})$ and $A^{\mathfrak{S}_\rho \mathbf{T}^x} = (A^{\mathbf{T}^x})^{\mathfrak{S}_\rho} = (A\zeta^{xa(\mathbf{T})})^{\mathfrak{S}_\rho}$ $= AD(S_\rho)\zeta^{xa(\mathbf{T})\,b(S_\rho)}$. Since $D(S_\nu S_\mu S_\nu{}^{-1} S_\mu{}^{-1}) = D(S_\rho)$ it follows that x is a solution of the congruence $xa(\mathbf{T})b(S_\rho) \equiv 1 \pmod{n}$. Thus the structure $G(K/F)$ is known.

Repeating the arguments of the proof of Theorem 19 and letting $c = 1$ we can assert

COROLLARY 7. *For a given normal extension $\mathfrak{W}$ of the residue class field $\mathfrak{F}$ and a given group $\Delta \supseteq \Gamma$ there exists at least*[25] *one normal extension L/F*

[25] The methods of [11] and the proof of Theorem 19 can be used to find the totality of all possible Galois groups for given finite W/F and Δ/Γ. A prerequisite for the solution of the "infinite" problem is the generalization of the theory of group extensions.

with $\mathfrak{L} = \mathfrak{W}$ and $\Gamma(L) = \Delta$ if and only if $\mathfrak{W}$ contains all those roots of unity whose orders occur as orders of the elements in Δ/Γ.

In order to obtain more specific results on the structure of the Galois groups we shall assume that the field F contains q elements. The unramified W_f is generated, according to Corollary 2, by a primitive $(q^f - 1)$-st root of unity Ω_f. The group $G(W_f/F)$ is cyclic and is generated by the Frobenius automorphism $\Omega_f \leftrightarrow \Omega_f^S = \Omega_f^q$.

Now let K/F be an extension of degree n with W_f as inertial field and $m_1, \cdots, m_r$ as a set of invariants of the factor group Δ/Γ, $\Delta = \Gamma(K)$. Then $n = f\prod_{i=1}^{r} m_i$. As before we have $K = K_1 \bigcup \cdots \bigcup K_r$ where $K_i = W([a_iC_iu_i]^{1/m_i})$. Since W_f is generated by Ω_f we may take the cyclic group $\{\Omega_f\}$ as a multiplicative set of representatives for the group $\mathfrak{W}^*$. Without loss of generality we may therefore suppose, by Lemma 1, $C_iu_i = \Omega_f^{h_i}$.

THEOREM 20. *Let K be a finite extension of F with maximal unramified subfield W_f/F and let $m_1, \cdots, m_r$ be the invariants of $\Gamma(K)/\Gamma$. Suppose $K = W([a_1C_1]^{1/m_1}, \cdots, [a_rC_r]^{1/m_r})$ where $C_i = \Omega_f^{h_i}$ and $\{(1/m_1)\phi(a_1), \cdots, (1/m_r)\phi(a_r), \Gamma\} = \Gamma(K)$, then K/F is normal if and only if* (i) *W_f contains all m_i-th roots of unity and* (ii) *$(m_i, q^f - 1) \mid h_i(q-1)$, $n = 1, 2, \cdots, r$.*

Proof. We rephrase the necessary and sufficient conditions which were obtained in Theorem 19. We have $C_i^{S-1} = D_i(S)^{m_i}$, $D_i(S)$ in W_f. Since C_i is a root of unity, we must have $D_i(S) = \Omega_f^{g_i}$. Consequently $\Omega_f^{h_i(S-1)} \Omega_f^{h_i(q-1)} = \Omega_f^{g_im_i}$ or $m_ig_i \equiv h_i(q-1) \pmod{q^f - 1}$. For given integers m_i the necessary and sufficient condition for the existence of g_i is given by $(m_i, q^f - 1) \mid h_i(q-1)$. The last relation imposes, for given Δ and W_f, conditions on the h_i. Hence K/F is normal if and only if $m_ig_i \equiv h_i(q^j - 1) \pmod{q^f - 1}$ for $j = 0, \cdots, f-1$. This set of conditions is equivalent to $(m_i, q^f - 1) \mid h_i(q-1)$. We obtain then m_i distinct solutions g_i which are in 1-1 correspondence with the m_i extensions of S to K_i

Employing the same notation as before we shall prove

THEOREM 21. *The Galois group of a normal extension K/F is generated by $r+1$ elements, $s, t_1, \cdots, t_r$ with the relations*

$$s^f = \prod_{i=1}^{r} t_i^{h_i}; \quad st_is^{-1} = t_i^q \quad \text{and} \quad t_i^{m_i} = 1, \quad \text{for} \quad i = 1, \cdots, r.$$

Proof.[26] We first determine the structure of the Galois group G_i of a typical field $K_i = W([a_iC_i]^{1/m_i})$ relative to F. By the Hilbert theory G_i

[26] An alternate, more explicit, proof may be given by using the formulas developed in the preceding proof.

has the following structure. The inertial field W/F belongs to the cyclic inertial group which is generated by an element of order m_i, say $\mathfrak{T}_i$. The factor group $G_i/\{\mathfrak{T}_i\}$ is isomorphic to $G(W/F) = \{S\}$ of order f. Let $\mathfrak{Z}_i$ be a representative of the coset S. By assumption on $\mathfrak{Z}_i$ we have $\mathfrak{Z}_i{}^f = \mathfrak{T}_i{}^{x_i}$ and $\mathfrak{Z}_i\mathfrak{T}_i\mathfrak{Z}_i{}^{-1} = \mathfrak{T}_i{}^{y_i}$. It remains to evaluate x_i, y_i. To find y_i we determine the effect of $\mathfrak{Z}_i\mathfrak{T}_i$ and $\mathfrak{T}_i{}^{y_i}\mathfrak{Z}_i$ on the elements of the field K_i/F. A simple computation proves $y_i = q$. Applying $\mathfrak{Z}_i{}^f$ to A_i we find $x_i = h_i$. We observe, as a consequence of above relations, that each g_i in G_i has a unique representation $\mathfrak{Z}_i{}^{a_i}\mathfrak{T}_i{}^{b_i}$.

Now let J_i be fixed isomorphisms of the factor groups $G_i/\{\mathfrak{T}_i\}$ upon $\{S\}$ such that the element $\mathfrak{Z}_i$ in G_i is mapped upon the automorphism S with $\Omega_f{}^S = \Omega_f{}^q$. Then $K = K_1 \cup \cdots \cup K_r$ implies, by a theorem of the Galois theory,[27] that the group $G(K/F)$ is isomorphic to a subgroup of the direct product $G_1 \times \cdots \times G_r$. The elements g of G are exactly those vectors $[g_1, \cdots, g_r] = [\mathfrak{Z}_1{}^{a_1}\mathfrak{T}_1{}^{b_1}, \cdots, \mathfrak{Z}_r{}^{a_r}\mathfrak{T}_r{}^{b_r}]$ whose components belong to the same cosets of $G_i/\{\mathfrak{T}_i\}$ by means of the isomorphisms J_i. Therefore we necessarily have $a_1 = \cdots = a_r = a$, i. e., $g = [\mathfrak{Z}_1{}^a\mathfrak{T}_1{}^{b_1}, \cdots, \mathfrak{Z}_r{}^a\mathfrak{T}_r{}^{b_r}]$. Let 1_i denote the unit of the group G_i. We define $t_i = [1_1, \cdots, 1_{i-1}, \mathfrak{T}_i, 1_{i+1}, \cdots, 1_r]$, $i = 1, \cdots, r$, and $s = [\mathfrak{Z}_1, \cdots, \mathfrak{Z}_r]$. These $r+1$ elements lie in $G(K/F)$. Since they lie in the direct product $G_1 \times \cdots \times G_r$ we have $t_it_j = t_jt_i$ of every pair i, j. The general element g of G can then be expressed as $s^a t_1{}^{b_1} \cdots t_r{}^{b_r}$ and the relations for the groups G_i imply [28] $s^f = \prod_{i=1}^{r} t_i{}^{h_i}$, $st_is^{-1} = t_i{}^q$, $t_i{}^{m_i} = 1$.

Let Γ'_∞ be the extension $(\prod_{p \neq \chi} p^\infty)^{-1}$ as determined by Theorem 1. Suppose that $\{a_{\nu_\mu}\}$ is a set of elements in F such that $(1/n_\nu)\phi(a_{\nu_\mu})$ exhausts the set of all elements of order n_ν in Γ'_∞/Γ, $n_\nu \mid \prod_{p \neq \chi} p^\infty$. We consider over F the unramified extension $W_\infty = \bigcup_n W_n$ where n runs over the set of all integers prime to χ. Then the Galois group $G(W_\infty/F)$ is isomorphic to the completion of the additive group of integers with respect to the topology determined by the principal ideals (n) as neighborhoods of zero. Let S be a generator of $G(W_\infty/F)$ such that $\Omega_n{}^S = \Omega_n{}^q$ for each primitive root of unity Ω_n of W_n. Next adjoin to W_∞ the set of radicals $a_{\nu\mu}{}^{1/n_\nu}$, $a_{\nu\mu}$ in $\{a_{\nu\mu}\}$, n_ν variable. We then obtain a field $K_\infty = W_\infty(\{a_{\nu\mu}{}^{1/n_\nu}\})$. This field is normal over F for it is approximated by the fields $W_{n\nu}(\{a_{\nu\mu}{}^{1/n_\nu}\})$ which are normal by Theorem 19. The group $G(K_\infty/W_\infty)$ is, by construction and Theorem 13,

[27] See [10].

[28] We emphasize that the elements $t_1, \cdots, t_r$ generate the abelian inertial group of K/F. The automorphism s is an extension of the generator S of $G(W/F)$.

equal to the abelian inertial group. Suppose that $\mathfrak{S}$ is an extension of S to $G(K_\infty/F)$.

THEOREM [20] 22. *Each element* $\mathfrak{T}$ *of* $G(K_\infty/W_\infty)$ *generates a cyclic invariant subgroup and* $\mathfrak{S}\mathfrak{T}\mathfrak{S}^{-1} = \mathfrak{T}^q$ *for a suitable extension* $\mathfrak{S}$ *of the automorphism* S *of* W_∞/F.

Proof. Let $\{\mathfrak{T}\}$ be the closed subgroup generated by $\mathfrak{T}$ in $G(K_\infty/W_\infty)$. The group $\{\mathfrak{T}\}$ is necessarily an ideal cyclic group for K_∞ is an infinite extension over the invariant field $M_{\mathfrak{T}} \supset W_\infty$ for $\{\mathfrak{T}\}$. For $[K_\infty : M_{\mathfrak{T}}] < \infty$ would imply $K_\infty = M_{\mathfrak{T}}(a^{1/n})$, a in $M_{\mathfrak{T}}$, such that the equation $n\xi = \phi(a)$ has no solution ξ in $\Gamma(M_{\mathfrak{T}})$. Hence the equations $n^i\xi_i = \phi(a)$, $i = 2, \cdots$, have no solutions ξ_i in $\Gamma(M_{\mathfrak{T}})$, contrary to the hypothesis on $K_\infty/M_{\mathfrak{T}}$. The group $\{\mathfrak{T}\}$ is an invariant subgroup for the field $M_{\mathfrak{T}}$ is generated (over W_∞) by radicals $\{a_{\nu'\mu'}{}^{1/n_{\nu'}}\} \subset \{a_{\nu\mu}{}^{1/n}\}$. Now let $L = W_\infty(b_i{}^{1/n_i})$ where K_∞ $M_{\mathfrak{T}}(b_i{}^{1/n_i})$. This field is normal over F since it is the unique extension over W_∞/F with $\{(1/n_i)\phi(b_i), \Gamma\}$ as value group. Therefore $G(L/W_\infty)$ is isomorphic to $G(K_\infty/M_{\mathfrak{T}}) = \{\mathfrak{T}\}$. The relation $L \cap M_{\mathfrak{T}} = W_\infty$ implies that $G(K_\infty/W_\infty)$ is the direct product of $G(L/W_\infty)$ and $G(M_{\mathfrak{T}}/W_\infty)$. Moreover, $G(K_\infty/F)$ is isomorphic to a subset of the direct product of $G(L/F)$ and $G(M/F)$. Next let $\mathfrak{S}'$ be the automorphism of $G(L/F)$ induced by $\mathfrak{S}$ and similarly let $\mathfrak{T}'$ be the automorphism induced by $\mathfrak{T}$. Again $\{\mathfrak{T}'\} \cong \{\mathfrak{T}\}$ is an invariant subgroup of $G(L/F)$. If $\mathfrak{S}\mathfrak{T}\mathfrak{S}^{-1} = \mathfrak{T}^x$ then also $\mathfrak{S}'\mathfrak{T}'\mathfrak{S}'^{-1} = \mathfrak{T}'^x$ by construction of L. Now let W_i be the unramified extension of F which contains the m_i-th roots of unity. Then the fields $L_i = W_i(b_i{}^{1/m_i})$ are normal over F, by Theorem 20, and $\mathfrak{S}_i\mathfrak{T}_i\mathfrak{S}_i{}^{-1} = \mathfrak{T}_i{}^q$ for suitable automorphisms $\mathfrak{S}_i$, $\mathfrak{T}_i$ of L_i/F. Letting $\bigcup_i L_i = L' \subseteq L$ the sequences $\{\mathfrak{S}_i\}$, $\{\mathfrak{T}_i\}$ determine elements $\mathfrak{S}''$, $\mathfrak{T}''$ of $G(L'/F)$ with $\mathfrak{S}''\mathfrak{T}''\mathfrak{S}''^{-1} = \mathfrak{T}''^q$. Therefore, on extending L' to L, $\mathfrak{S}'\mathfrak{T}'\mathfrak{S}'^{-1} = \mathfrak{T}^q$. We let then $\mathfrak{S}$ be an extension of $\mathfrak{S}'$.

THEOREM 23. *The field* K_∞/F *contains all extensions* L/F *whose degrees are prime to* χ.

Proof. When L/F is an infinite extension we say that its degree is prime to χ if each element of L is the root of an equation in F whose degree is prime to χ. Let W be the maximal unramified subfield of L/F. Then $W \subseteq W_\infty$. Suppose that $L = W(\{A_{\nu\mu}{}^{1/n_\nu}\})$, $A_{\nu\mu}$ in W with $\{(1/n_\nu)\phi(A_{\mu\nu}), \Gamma\} = \Gamma(L)$. Since W_∞ admits no unramified extensions we have $W_\infty \cup L$ $W_\infty(\{A_{\nu\mu}{}^{1/n_\nu}\}) = W_\infty(\{a_{\nu\mu}{}^{1/n_\nu}\})$ where $a_{\nu\mu}$ in F and $\phi(a_{\nu\mu}) = \phi(A_{\nu\mu})$. Consequently $K_\infty \supseteq W_\infty(\{a_{\nu\mu}{}^{1/n_\nu}\}) = W_\infty \cup L \supseteq L$.

[20] A similar theorem may be proved for arbitrary normal fields L/F.

We next use a modification of the preceding arguments for the discussion of the normal extensions of degree p^h over F. In case that F does not contain a primitive p-th root of unity or $\Gamma = p\Gamma$, Theorem 16 states that the extensions K must be cyclotomic. We thus shall suppose that F contains all p^μ-th roots of unity but not a primitive $p^{\mu+1}$-th root of unity, $\mu \geqq 1$. Moreover, we shall suppose, for the sake of simplicity, that $[\Gamma : p\Gamma] = p^h$ for the given prime p.

Let $a_1, \cdots, a_h$ be a set of elements in F so that the residue classes $\phi(a_i)$ mod Γ are independent generators for $\Gamma/p\Gamma$. Furthermore let W_j be the unramified extension of degree p^j over F. We define $K_j = W_j((a_1)^{1/p^j}, \cdots, (a_h)^{1/p^j})$. Then $K_{j-1} \subset K_j$ and $[K_j : F] = p^{j(h+1)}$ where we set $K_0 = F$.

THEOREM 24. *The field $F^{(p)} = \bigcup_j K_j$ is the smallest universal field for all p-extensions over F. The Galois group $G(F^{(p)}/F)$ is generated by $h+1$ elements $\mathfrak{S}, \mathfrak{T}_1, \cdots, \mathfrak{T}_h$ with the relations $\mathfrak{S}\mathfrak{T}_i\mathfrak{S}^{-1} = \mathfrak{T}_i^q$ for $i = 1, \cdots, h$. Each element of $G(F^{(p)}/F)$ has a unique representation $\mathfrak{S}^\alpha \mathfrak{T}_1^{\beta_1} \cdots \mathfrak{T}^{\beta_h}$ where $\alpha, \beta_1, \cdots, \beta_h$ are p-adic integers.*

Proof. The join $F^{(p)} = \bigcup_j K_j$ can be expressed as $W^{(p)}((a_i)^{1/p^j})$ where $W^{(p)} = \bigcup_j W_j$ and $i = 1, \cdots, h$, $j = 1, 2, \cdots$. Since the fields $W_j(a_i^{1/p^j})$ are normal over F, the field $F^{(p)}$ is normal, too. The field $F^{(p)}$ contains all normal p-extensions. By the usual device of approximation it suffices to prove $K \subset F^{(p)}$ where K/F is an arbitrary finite p-extension. Suppose then that $F \subset K^{(1)} \subset \cdots \subset K^{(k)} = K$ is a chain of relatively cyclic fields $K^{(\nu)}/K^{(\nu+1)}$ of degree p. By construction K_1 is the maximal abelian extension of exponent p over F, so that $K^{(1)} \subset K_1$. Therefore $[K \cup K_1 : K_1] < [K : F]$. Suppose that $K^{(\nu_1)} \subseteqq K_1$, $K^{(\nu_1+1)} \not\subset K_1$, $\nu_1 > 1$. Then K_2 absorbs $K^{(\nu_1+1)}$ and possibly some of its successors in the chain for K/F for K_2 is the maximal abelian extension of exponent p over K_1. Hence $K \subseteqq K_j$, for sufficiently large j, since $[K : F]$ is finite. Finally, $F^{(p)}$ is minimal by construction.

In order to determine the structure of $G(F^{(p)}/F)$ we let $F^{(p)} = \bigcup_i [\bigcup_j W_j(a_i^{1/p^j})] = \bigcup_i W^{(p)}(\{a_i^{1/p^j}\})$. Since $G_i = G(W^{(p)}(\{a_i^{1/p^j}\})/F) \lim_j G(W_j{}^i(a_i^{1/p^j}\mathfrak{Q}/F)$ we can find elements $\mathfrak{S}'_i$, $\mathfrak{T}'_i$ in G_i such that $\mathfrak{S}'_i\mathfrak{T}'_i\mathfrak{S}'^{-1}_i = \mathfrak{T}'^q_i$. Moreover, each element of G_i has a unique representation as $\mathfrak{S}'^{\sigma_i}_i \mathfrak{T}'^{\tau_i}_i$ where the σ_i, τ_i are p-adic integers. These statements are implied [30] by Theorem 21 and the fact that both $G(W^{(p)}/F)$ and $G(W^{(p)}(\{a_i^{1/p^j}\})/W^{(p)})$ are isomorphic to the additive group of p-adic integers. To complete the proof we observe that $G(F^{(p)}/F)$ is a subgroup of the direct product $G_1 \times \cdots \times G_h$. A slight generalization of the arguments

[30] Letting $G_i = \lim_j G_{ji}$ where $G_{ji} = G(W_j(a_i^{1/p^j})/F)$.

used in the proof of Theorem 21 establishes the remaining assertions of the theorem.

The same methods as above may be employed to prove

THEOREM 25. *If* $[\Gamma: p_\nu\Gamma] = p_\nu^{N(\nu)}$ *for all rational primes* $p_\nu \neq \chi$ *and* $\max_\nu N(\nu) = N < \infty$, *then the Galois group* $G(K_\infty/F)$ *can be generated by* $N+1$ *elements* $\mathfrak{S}, \mathfrak{T}_1, \cdots, \mathfrak{T}_N$ *with the relations* $\mathfrak{S}\mathfrak{T}_j\mathfrak{S}^{-1} = \mathfrak{T}_j^q$ *for* $j = 1, \cdots, N$. *The elements* $\mathfrak{T}_1, \cdots, \mathfrak{T}_N$ *generate the (abelian) inertial group of* K_∞/F. *Each element of* $G(K_\infty/F)$ *has a unique representation* $\mathfrak{S}^\alpha\mathfrak{T}_1^{\beta_1}\cdots\mathfrak{T}_N^{\beta_N}$ *where the elements* $\alpha, \beta_1, \cdots, \beta_N$ *lie in suitable completions*[31] *of the additive group of integers.*

THEOREM 26. *A finite group* G *of order* p^n *is realizable as the Galois group of a normal extension* K/F *if and only if* G *is the homomorphic map of the infinite discrete group* G_p *with the generators* $\mathfrak{S}, \mathfrak{T}_1, \cdots, \mathfrak{T}_h$ *and the relations* $\mathfrak{S}\mathfrak{T}_i\mathfrak{S}^{-1} = \mathfrak{T}_i^q$ *for* $i = 1, \cdots, h$.

Proof. The group G_p of the theorem is an everywhere dense subgroup of $G(F^{(p)}/F)$ for the elements $\mathfrak{S}, \mathfrak{T}_1, \cdots, \mathfrak{T}_h$ of $G(F^{(p)}/F)$ determine, by the homomorphism $G(F^{(p)}/F) \to G(K_j/F)$, the generators of each group $G(K_j/F)$. Hence the group G_p can be interpreted, by Theorem 3, as an everywhere dense subgroup of $G(F^{(p)}/F)$. Now let K/F be an arbitrary finite p-extension. Since $K \subset F^{(p)}$ it follows, by Theorem 3, that $G(K/F)$ is a homomorphic map of G_p. Conversely, let S be a homomorphic map of G_p, say $G \cong G_p/S$. Suppose that G_p is imbedded in $G(F^{(p)}/F)$. Then S determines, by Theorem 3, a subfield K of $F^{(p)}$ with $G(K/F) \cong G$.

Our methods can be used to prove

THEOREM 27. *A group* G *can be realized as the Galois group of a normal extension* L/F *if and only if* (i) G *has an ideal cyclic factor group* G_ϕ *such that thai the kernel* H *of the homomorphism* $G \to G_\phi$ *is a locally compact group satisfying the second axiom of countability and is the character group of some extension* $\Delta \supseteq \Gamma$ *and* $\mathfrak{S}\mathfrak{T}\mathfrak{S}^{-1} = \mathfrak{T}^q$ *for each* $\mathfrak{T}$ *in* H *and for an extension* $\mathfrak{S}$ *of the Frobenius automorphism* S *of a suitable unramified* W/F, *and* (ii) *the field* W (*belonging to* G_ϕ) *contains all those roots of unity whose orders occur for the elements of* Δ/Γ.

THE UNIVERSITY OF CHICAGO.

---

[31] The elements α lie in the completion of the additive group of integers Γ_0 with respect to the principal ideals (n), $n \not\equiv 0 \pmod{\chi}$, as a fundamental system of neighborhoods of zero in Γ_0. To determine the groups $\{\beta_j\}$ we let $M_j = \bigcup_j L_{\nu j}$ where $L_{\nu j} = W_\infty(\{a_{\nu j}^{1/n_j}\})$ and the $a_{\nu j}$ are in F so that $\phi(a_{\nu j})$ are representives for the generators of $\Gamma/p_\nu\Gamma$, p_ν rational primes. Then $G(M_j/W_\infty) \cong \{\beta_j\}$ with a topology similar to that of $\{\alpha\}$.

BIBLIOGRAPHY

1. A. A. Albert, *Modern Higher Alegbra*, Chicago (1937).

2. A. A. Albert, "On p-adic fields and rational division algebras." *Annals of Mathematics*, vol. 41 (1940), pp. 674-93.

3. R. Baer, "Abelian fields and duality of abelian groups," *American Journal of Mathematics*, vol. 59 (1937), pp. 869-88.

4. M. Deuring, "Verzweigungstheorie bewerteter Körper," *Mathematische Annalen*, vol. 105 (1931), pp. 277-307.

5. J. Herbrand, "Théorie arithmétique des corps de nombres de degré infini. II. Extensions algébriques de degré infini," *Mathematische Annalen*, vol. 108 (1933), pp. 699-717.

6. W. Krull, "Galoissche Theorie unendlichen algebraischen Erweiterungen," *Mathematische Annalen*, vol. 100 (1928), pp. 687-98.

7. W. Krull, "Galoissche Theorie bewerteter Körper," *Sitzungsber. der Akad. München, Math.-Natur. Klasse* (1930-31), pp. 225-38.

8. W. Krull, "Allgemeine Bewertungstheorie," *Jour. für d. reine und ang. Math.*, vol. 167 (1932), pp. 169-96.

9. S. MacLane and O. F. G. Schilling, "Zero-dimensional branches of rank one on algebraic varieties," *Annals of Mathematics*, vol. 61 (1939), pp. 883-96.

10. S. MacLane and O. F. G. Schilling, "A formula for the direct product of cross product algebras," *Bulletin of the American Mathematical Society*, vol. 49 (1942), pp. 108-114.

11. S. MacLane and O. F. G. Schilling, "A general Kummer theory for function fields," *Duke Mathematical Journal*, vol. 9 (1942), pp. 125-167.

12. M. Moriya, "Theorie der algebraischen Zahlkörper unendlichen Grades," *Jour. of the Fac. of Sc.*, The Hokkaido Imperial University, vol. 3 (1935), pp. 107-90.

13. M. Moriya, "Galoissche Theorie der algebraischen Zahlkörper unendlichen Grades, *Jour. of the Fac. of Sc.*, The Hokkaido Imperial University, vol. 4 (1936), pp. 67-120.

14. O. Ore, "Über den Zusammenhang zwischen den definierenden Gleichungen und der Idealtheorie in algebraischen Körpern I," *Mathematische Annalen*, vol. 96 (1926), pp. 313-52, II, *Mathematische Annalen*, vol. 97 (1927), pp. 569-98.

15. A. Ostrowski, "Einige Fragen der allgemeinen Körpertheorie," *Jour. für. d. reine und ang Math.*, vol. 143 (1913), pp. 255-84.

16. A. Ostrowski, "Untersuchungen zur arithmetischen Theorie der Körper," *Mathematische Zeitschrift*, vol. 39 (1934), pp. 269-404.

17. L. Pontrajagin, *Topological Groups*, Princeton (1939).

18. O. F. G. Schilling, "A generalization of local class field theory, *American Journal of Mathematics*, vol. 60 (1938), pp. 667-704.

19. O. F. G. Schilling, "Arithmetic in fields of formal power series in several variables," *Annals of Mathematics*, vol. 38 (1937), pp. 553-76.

20. O. F. G. Schilling, "Regular normal extensions over complete fields," *Transactions of the American Mathematical Society*, vol. 47 (1940), pp. 440-54.

21. N. E. Steenrod, "Universal homology groups," *American Journal of Mathematics*, vol. 58 (1936), pp. 661-701.

22. E. Steinitz, *Algebraische Theorie der Körper*, Berlin (1930).

23. H. Zassenhaus, *Lehrbuch der Gruppentheorie*, Leipzig (1937).

REPRESENTATION OF SUBHARMONIC FUNCTIONS IN THE NEIGHBORHOOD OF A POINT.*

By W. R. TRANSUE.

It was F. Riesz [1] who first took up the problem of the representation of a subharmonic function in its domain of subharmonicity. He found that a subharmonic function is not in general representable as a logarithmic potential in any domain, D, in which it is subharmonic, but only in a domain D' contained, together with its boundary, in D. By "representable as a logarithmic potential in D'" is meant [2] that there exists a positive mass-distribution $\mu(e)$ such that

$$u(P) = \int_{D'} \log PQ \, d\mu(e_Q) + h(P); \qquad P \text{ in } D'$$

where Stieltjes-Radon integration is indicated and $h(P)$ is harmonic in D'.

To show this Riesz constructed the following example. Consider the function

$$u(P) = \sum_{k=1}^{\infty} \log \frac{PQ_k}{PQ_{-k}}$$

where Q_k and Q_{-k} are points on the X axis whose abscissas are $1/k$ and $-1/k$ respectively. This function is subharmonic in the interior of the circle with center at $(1, 0)$ and radius 1, but any attempt to represent it as a logarithmic potential leads to a function which is everywhere infinite.

Reisz further found [3] that, using generalized positive mass-distributions [4] and replacing the function $\log PQ$ by $G(P, Q)$, the Green's function for D with pole at Q, u could be represented by

$$u(P) = -\int_D G(P, Q) \, d\mu(e_Q) + h(P); \qquad P \text{ in } D$$

where $h(P)$ is harmonic in D, for any domain D in which u is subharmonic,

* This paper is part of a dissertation prepared at Lehigh University under the direction of Professor G. E. Raynor; Received October 30, 1941.

[1] F. Riesz, "Sur les fonctions subharmoniques et leur rapport à la théorie du potentiel, Part 1," *Acta Mathematica*, vol. 48 (1926), pp. 329-343.

[2] See Radó, "Subharmonic Functions," *Ergebnisse der Mathematik*, Berlin, 1937, Chap. IV.

[3] F. Riesz, "Sur les fonctions subharmoniques et leur rapport à la théorie du potentiel, Part II," *Acta Mathematica*, vol. 54 (1930), pp. 321-360.

[4] Rado, *loc. cit.*, 6. 13.

provided only that u possess an harmonic majorant in D. Since in the example given above the constant 0 is an harmonic majorant for u in the circle, it follows that it is representable there by use of Green's function while it was not as a logarithmic potential. Replacing $\log PQ$ by $G(P,Q)$ led from an integral which did not converge to one which did.

The question now naturally arises as to whether, by replacing $G(P,Q)$ by some other familiar function of P and Q, it is not possible to obtain a representation for subharmonic functions in domains in which they do not possess an harmonic majorant, and hence for which the representation by means of Green's function fails. I shall endeavor to show that this is possible under certain conditions.

For this purpose we shall consider functions subharmonic in the whole plane but not at infinity, that is, functions which, when transformed by an inversion in a circle about the origin, go over into functions subharmonic in any neighborhood of the origin but not in any domain containing the origin itself. The condition that such a function possess an harmonic majorant in the whole plane can be obtained by means of a transformation by inversion from the results of Brelot [5] for the case of a function subharmonic in the neighborhood of the origin. These results when transformed show that if u' is subharmonic in the domain exterior to a closed contour, B, but not at infinity, a necessary and sufficient condition that u' possess an harmonic majorant in this infinite exterior domain is that $\lim_{r\to\infty} \frac{M_r u'}{\log r} \equiv \lambda_m(\infty)$ be finite, where $M_r u'$ denotes the mean of u' on the circumference of a circle about the origin of radius r.

Privaloff [6] has shown that

$$M_r u = \int_0^r \frac{n(\rho,u) - n(0,u)}{\rho}\, d\rho + n(0,u)\log r + C$$

where $n(\rho,u)$ is the mass in the interior of the circle about the origin with radius ρ in the representation of u as a potential, $n(0,u)$ is the mass at the origin, and C is a constant. We see, then, that if the mass inside a circle of radius r in the representation of u as a potential does not increase as r increases, u may be represented by means of the function $\log PQ$; if the mass is infinite but does not increase so rapidly that $M_r u$ increases faster than $\log r$, then u may be represented by means of the function $G(P,Q)$; if the mass increases

[5] M. Brelot, "Étude des fonctions sousharmoniques au voisinage d'un point," *Actualités Scientifiques et Industrielles*, No. 139, Paris, 1934.

[6] I. I. Privaloff, "Sur la généralisation d'une formule de Jensen, I, *Bull. Acad. Sci.*, URSS (1935), pp. 837-847.

faster than this neither of these representations is available and a new method of representation must be found.

To show that such functions are not unusual let us recall that integral transcendental functions of finite order are those for which

$$\log |f(z)| < |z|^{\alpha}$$

where α is a constant. Since $\log |f(z)|$ is a subharmonic function we may obtain, without leaving the class of functions of finite order, examples of subharmonic functions which are not representable by means of Green's function.

Let us note that in obtaining Green's function an harmonic function is added to $\log PQ$ and that the resulting function represents u where $\log PQ$ fails. It seems natural, then, to look for an harmonic function to add to $\log PQ$ which will represent functions where $G(P, Q)$ fails. Let us consider the function

$$\begin{aligned} E_p(P,Q) \equiv E_p(x,y,\xi,\eta) &\equiv \log PQ - \log OQ \\ &\quad + \mathcal{R}\left[\frac{P}{Q} + \frac{1}{2}\left(\frac{P}{Q}\right)^2 + \cdots + \frac{1}{p}\left(\frac{P}{Q}\right)^p\right] \\ &\equiv \log \sqrt{(x-\xi)^2 + (y-\eta)^2} - \log \sqrt{\xi^2+\eta^2} \\ &\quad + \mathcal{R}\left[\frac{(x+iy)}{(\xi+i\eta)} + \frac{1}{p}\frac{(x+iy)^2}{(\xi+i\eta)^p} + \cdots + \frac{1}{p}\frac{(x+iy)^p}{(\xi+i\eta)^p}\right] \\ &\equiv \log PQ + H_p(P,Q) \end{aligned}$$

where (x, y) are the coördinates of the point P and (ξ, η) are those of Q, and $\mathcal{R}$ denotes the real part of the complex number which follows it. For $\xi^2 + \eta^2 \neq 0$, $H_p(P, Q) = E_p(P, Q) - \log PQ$ is an harmonic function of Q for fixed P since $\left[\frac{P}{Q} + \frac{1}{2}\left(\frac{P}{Q}\right)^2 + \cdots + \frac{1}{p}\left(\frac{P}{Q}\right)^p\right]$ is an analytic function of Q, and, for fixed Q, $H_p(P, Q)$ is an harmonic function of P in the whole plane.

Now suppose that $u(P)$ is subharmonic in the whole plane (but not at ∞). Such a function may, as remarked above, be represented inside any circle about the origin as a logarithmic potential plus an harmonic function. Let us turn our attention to the problem of representing it outside such a circle of radius b. Let C_R denote the circle about the origin with radius R and let ${}_bC_R$ designate the domain interior to C_R and exterior to $C_b (R > b)$. Then there exists in the plane a generalized positive mass-distribution $\mu(e)$ such that

$$u(P) = \int_{{}_bC_R} \log PQ \, d\mu(e_Q) + H(P); \qquad P \text{ in } {}_bC_R$$

where $H(P)$ is harmonic in ${}_bC_R$. Then, since the integrand differs from $\log PQ$ by an harmonic function of Q, the integral

$$v_R^{(p)}(P) = \int_{{}_bC_R} E_b(P,Q)\, d\mu(e_Q)$$

exists almost everywhere in ${}_bC_R$ and since

$$v_R^{(p)}(P) = \int_{{}_bC_R} \log PQ\, d\mu(e_Q) + \int_{{}_bC_R} H_p(P,Q)\, d\mu(e_Q)$$

$v_R^{(p)}(P)$ differs from $u(P)$ in ${}_bC_R$ by a function harmonic there and hence $v_R^{(p)}(P)$ is subharmonic in ${}_bC_R$.

Consider the integral

$$I_R^{(p)} = \int_{{}_bC_R} \frac{1}{|\,OQ\,|^{p+1}}\, d\mu(e_Q).$$

Let us show that if, as $R \to \infty$, $I_R^{(p)}$ approaches a limit $I^{(p)}$, then $v_R^{(p)}(P)$ will approach a limit $v^{(p)}(P)$ uniformly on any closed set in the plane. Let S be a closed set in the plane enclosed by the circle C_r. It is sufficient to show that for any point P in S and for any $\epsilon > 0$, there exists a ρ such that for $R > \rho$ and $t > 0$, $|\, v_R^{(p)}(P) - v_{R+t}^{(p)}(P)\,| < \epsilon$. Now

$$v_R^{(p)}(P) - v_{R+t}^{(p)}(P) = \int_{{}_RC_{R+t}} E_p(P,Q)\, d\mu(e_Q).$$

Blumenthal [7] has shown that

$$E_p(P,Q) > -\left|\frac{OP}{OQ}\right|^{p+1} \quad \text{and} \quad E_p(P,Q) < \left|\frac{OP}{OQ}\right|^{p+1}$$

provided that $\left|\frac{OP}{OQ}\right| < 1 - \frac{1}{p+1}$. Hence $|\, E_p(P,Q)\,| < \left|\frac{OP}{OQ}\right|^{p+1}$ if $\left|\frac{OP}{OQ}\right| < 1 - \frac{1}{p+1}$. It follows that

$$|\, v_R^{(p)}(P) - v_{R+t}^{(p)}(P)\,| = \left|\int_{{}_RC_{R+t}} E_p(P,Q)\, d\mu(e_Q)\right| \leqq \int_{{}_RC_{R+t}} |\, E_p(P,Q)\,|\, d\mu(e_Q)$$

$$< \int_{{}_RC_{R+t}} \left|\frac{OP}{OQ}\right|^{p+1} d\mu(e_Q) = |\, OP\,|^{p+1} \int_{{}_RC_{R+t}} \frac{1}{|\,OQ\,|^{p+1}}\, d\mu(e_Q)$$

[7] O. Blumenthal, *Principes de la Théorie des Fonctions Entières d'Ordre Infini*, Paris (1910), p. 51.

where in making use of the inequality $|E_p(P,Q)| < \left|\frac{OP}{OQ}\right|^{p+1}$ we of course assume that the condition $\left|\frac{OP}{OQ}\right| < 1 - \frac{1}{p+1}$ under which this inequality holds is fulfilled. This can be done for any p by taking R sufficiently large since $|OP| \leqq r$ and $|OQ| \geqq R$ and thus $\left|\frac{OP}{OQ}\right| \leqq \frac{r}{R}$. Now, if $I_R^{(p)}$ converges, the last integral above may be made arbitrarily small by choosing R sufficiently large. In case $v_R^{(p)}(P)$ converges we define

$$v^{(p)}(P) = \lim_{R\to\infty} v_R^{(p)}(P) \equiv \int_{bC_\infty} E_p(P,Q)\, d\mu(e_Q).$$

Since

$$u(P) - v_R^{(p)}(P) = H_R^{(p)}(P)$$

where $H_R^{(p)}(P)$ is harmonic in ${}_bC_R$, and since $v_R^{(p)}(P)$ converges uniformly to $v^{(p)}(P)$ on any closed set, $H_R^{(p)}(P)$ also converges uniformly on any closed set to a function $H^{(p)}(P)$. Because the functions $H_R^{(p)}(P)$ are equal to their means on any circle and because the convergence is uniform, the limit of the mean of $H_R^{(p)}$ is equal to the mean of $H^{(p)}(P)$. Therefore $H^{(p)}(P)$ is equal to its mean on any circle and, by the converse of Gauss' theorem, we have the fact that $H^{(p)}(P)$ is harmonic in the domain exterior to C_b. Hence

$$u(P) = \int_{bC_\infty} E_p(P,Q)\, d\mu(e_Q) + H^{(p)}(P)$$

and we have a representation for the function $u(P)$ using $E_p(P,Q)$. Thus we have the

THEOREM. *If u is subharmonic in the whole plane and $\mu(e)$ the corresponding generalized positive mass-distribution, and if* $\lim_{R\to\infty} \int_{bC_R} \frac{1}{|OQ|^{p+1}}\, d\mu(e_Q)$ *exists, where p is a positive integer and ${}_bC_R$ denotes the domain between the circles with center at the origin and radii b and R respectively, then*

$$u(P) = \lim_{R\to\infty} \int_{bC_R} E_p(P,Q)\, d\mu(e_Q) + H^{(p)}(P) \equiv \int_{bC_\infty} E_p(P,Q)\, d\mu(e_Q) + H^{(p)}(P)$$

for P outside the circle of radius b, where $E_p(P,Q) \equiv \log PQ - \log OQ + \Re\left[\frac{P}{Q} + \frac{1}{2}\left(\frac{P}{Q}\right)^2 + \cdots + \frac{1}{p}\left(\frac{P}{Q}\right)^p\right]$ *P and Q in the right member standing*

for the complex numbers corresponding to these two points. The function $H^{(p)}(P)$ *is harmonic outside the circle with radius* b.

It may be remarked that, while in the statement of the preceding theorem we considered functions subharmonic in the whole plane, the proof given will apply equally well to functions which are merely subharmonic in the domain exterior to some closed contour. In this case the circle C_p outside which the function is represented must lie entirely within this exterior domain.

Let us now consider a function u, subharmonic in the neighborhood of the origin. Then any one of three cases may occur.

1. u is bounded above in the neighborhood of the origin, 0, and can be made subharmonic there by defining $u(0) = \overline{\lim_{P \to 0}} u(P)$.

2. u is not bounded above in the neighborhood of 0 but $\lim_{r \to 0} \frac{M_r u(0)}{\log r}$ exists and hence u possesses an harmonic majorant in the neighborhood of 0. In this case u may be represented in the neighborhood of the origin using Green's function. This case has been discussed in detail by Brelot.[8]

3. u does not have an harmonic majorant in the neighborhood of 0. In this case u cannot be represented in the neighborhood of 0 either as a logarithmic potential or by means of Green's function.

However if we transform u by an inversion in the unit circle it goes over into a function subharmonic outside a sufficiently large circle. As we have seen, this function may be represented using the function $E_p(P, Q)$ provided the mass does not increase too rapidly. So, on transforming the function $E_p(P, Q)$ by an inversion, a representation of u in the neighborhood of the origin may be obtained in this case.

While the above theorem widens considerably the class of functions subharmonic in the whole plane (or, equivalently, in the neighborhood of a point) which are representable as the sum of an harmonic function and an integral in terms of known functions, it does not solve this problem completely. To treat the general problem of representing any function subharmonic in the whole plane by this method it will be necessary to allow the index p in the function $E_p(P, Q)$ to approach ∞ with R in defining the integral. It would seem that a treatment of this general case could be built up following Blumenthal's discussion of integral functions of infinite order.[9]

LEHIGH UNIVERSITY.

---

[8] M. Brelot, *loc. cit.*[5]

[9] O. Blumenthal, *loc. cit.*[7]

VARIATION OF THE GREEN FUNCTION AND THEORY OF THE p-VALUED FUNCTIONS.*

By MENAHEM SCHIFFER.

1. Introduction. Let Δ be a domain in the ζ-plane, which contains the point $\zeta = 0$ but not $\zeta = \infty$. We consider the family of all functions

$$f(\zeta) = \zeta + a_2\zeta^2 + a_3\zeta^3 + \cdots + a_n\zeta^n + \cdots, \tag{1}$$

regular and univalent in Δ; this family is known to be compact and therefore we may ask for the extremal functions $f_n(\zeta)$ in the family possessing the largest value of $|a_n|$.

Each extremal function $z = f_n(\zeta)$ maps Δ on a domain D_n in the z-plane, having the same type as Δ and possessing as boundary continua (if there are such at all) analytic curves. By means of the identity

$$f(\zeta)[1 - xf(\zeta)]^{-1} = \sum_{\nu=1}^{\infty} \{a_\nu + P_\nu(x)\}\zeta^\nu \tag{2}$$

we may attach to each function $f(\zeta)$ a set of polynomials $P_\nu(x)$; then the differential equations, satisfied by the boundary curves of D_n, can be written in the following form. Let $P_n(x)$ be the n-th polynomial, belonging to $f_n(\zeta)$, and let $z(t)$ be a parametric representation of a fixed boundary curve, then we get by appropriate choice of the parameter t [1]

$$\frac{1}{a_n}\left(\frac{z'(t)}{z(t)}\right)^2 P_n\left(\frac{1}{z(t)}\right) + 1 = 0. \tag{3}$$

In the special case where Δ is simply connected and, therefore, can be supposed to be the unit circle, it is conjectured that $|a_n| \leq n$ holds and that

$$b_n(\zeta) = \zeta(1-\zeta)^{-2} = \sum_{\nu=1}^{\infty} \nu\zeta^\nu \tag{4}$$

is, essentially, the only extremal function. Indeed, $b_n(\zeta)$ satisfies the conditions above and there arises the question: Do these determine the extremal function completely. In the equation (3) there appear the coefficients of the

* Received June 20, 1941; Revised August 21, 1942.

[1] M. Schiffer, "A method of variation within the family of simple functions," *Proceedings of the London Mathematical Society*, Ser. 2, Vol. 44 (1938), pp. 432-449. M. Schiffer, "On the coefficients of simple functions," *Proceedings of the London Mathematical Society*, Ser. 2, Vol. 44 (1938), pp. 450-452.

required function; we have, therefore, a difficult functional equation, the solution of which would be decisive for the coefficient problem in the theory of univalent functions.

The above-mentioned results were obtained by means of a method of variations which is based on theorems concerning point sets and conformal representation. In this paper we shall develop an alternative method, using only the most elementary theorems of potential theory, and obtain the same results as those previously obtained in the case of a simply connected domain Δ. In particular, we get in this way a new proof, based on potential theory, of the "$\frac{1}{4}$-theorem" of Koebe-Pick. In the case of multiple connectivity of Δ, the new method differs essentially from the former and is applicable to problems where this failed, but is unapplicable to questions where the old method was useful. The reason for this difference is the following: the old method of variations transforms D_n into a domain of comparison D^*_n of the same conformal type (i. e. D^*_n can be mapped conformally on D_n), while this does not hold for the new method. The latter is, therefore, preferable in problems of pure potential theory, while the former remains valuable in questions of conformal representation of multiply connected domains.

Finally, it will be shown that the new method of variations is also applicable to Riemann surfaces (R. S.) and can be used, therefore, in the theory of p-valued functions.

2. Definition and properties of the variations considered. Let us suppose, for the sake of generality, that we are dealing with a closed R. S. $\mathfrak{R}$, p-sheeted and of genus g. Let $q(z)$ be a uniform function on $\mathfrak{R}$, regular everywhere with the exception of the points $z_0, z_1, \cdots z_m$, which are supposed to be finite and not branch points. At each z_i let $q(z)$ possess a simple pole and the development

$$q(z) = a_i(z-z_i)^{-1} + b_i + c_i(z-z_i) + \cdots. \tag{5}$$

Then, the function

$$z^* = z + \rho q(z) \tag{6}$$

will be regular on $\mathfrak{R}$ with the exception of $z = \infty$ and $z = z_i$. Around each point z_i we describe a circle k_i with radius r, and r is chosen so small that no branch point and no point z_k ($\neq z_i$) lies in the interior $\tilde{k}_i$ of k_i. Set

$$m = \operatorname{Max} |q(z)|, \qquad |z-z_i| = \tfrac{1}{2}r, \qquad i = 0, 1, \cdots m. \tag{7}$$

If ρ satisfies the inequality

$$|\rho| < r/2m \tag{8}$$

then $z^*(z)$ is exactly p-valued over the domain $\mathfrak{R} - \sum_{i=0}^{m} \tilde{k}_i$. Indeed, let a be a point of the domain; then, according to (7) and (8), we have, for $|z - z_i| = r/2$, $|z - a| > |\rho q(z)|$; it follows, therefore, from the theorem of Rouché that $z^*(z)$ takes on the value a in $\mathfrak{R} - \sum_{i=0}^{m} k_i$ at most p times. On the other hand, each value in the neighborhood of ∞ is taken on exactly p times. Therefore $z^*(z)$ is p-valued in $\mathfrak{R} - \sum_{i=0}^{m} \tilde{k}_i$. The representation $z^*(z)$ transforms the circles k_i into analytic curves k^*_i without self-intersections, if only we have

$$(9) \qquad |\rho| < (\text{Max} \left| \frac{q(x) - q(y)}{x - y} \right|)^{-1}, \qquad x,y \subset k_i, \qquad i = 0, \cdots m,$$

a condition compatible with (8).

Let us summarize: The representation $z^*(z)$ transforms the domain $\mathfrak{R} - \sum_{i=0}^{m} \tilde{k}_i$, contained in the R. S. $\mathfrak{R}$, into a domain bounded by $(m+1)$ simple curves k^*_i and covering the z^*-plane p times at most. If we add to this domain the interiors $\tilde{k}^*_i$ of the curves k^*_i, we get a closed R. S. $\mathfrak{R}^*_{\rho q}$ with p sheets. $\mathfrak{R}^*_{\rho q}$ will be called the R. S. obtained from $\mathfrak{R}$ by means of the variation $V_{\rho q}$.

This variation can be obtained by continuous deformation of $\mathfrak{R}$; it does not alter, therefore, the genus of the surface, since this is a topological invariant. Thus, the variations $V_{\rho q}$ preserve the number of sheets and the genus of the R. S.

Let D be a domain on $\mathfrak{R}$. If no pole z_i of $q(z)$ is situated on the boundary of D, and if r is so small that no point of this boundary is situated in any $\tilde{k}_i$, then $V_{\rho q}$ determines in an unambiguous way a variation $D \to D^*_{\rho q}$. Therefore, a domain D on a closed p-sheeted R. S. $\mathfrak{R}$ with genus g is transformed by means of a variation $V_{\rho q}$ into a domain D^* on a R. S. $\mathfrak{R}^*$ of the same type.

It is evident that analogous variations can be defined on non-closed R. S.; for our purposes, however, it will not be necessary to do this.

3. The variation of the Green function. Let us consider a domain D, on a R. S. $\mathfrak{R}$ of the above type, which is bounded by analytic curves. Then the Green functions $g(z;x)$ are known to exist on D, behaving near $z = x_j$ as $-\log|z - x_j|$, harmonic elsewhere in D, and converging to zero when the argument approaches a point of the boundary of D. Let us submit D to a variation $V_{\rho q}$ as defined in **2**. For sufficiently small ρ we get a domain D^*

on a R. S. $\mathfrak{R}^*$ with the same properties; we also have new Green functions $g^*(z;x)$ defined for z and x in D^*. Our aim will be to compute $g^*(z;x)$, from $g(z;x)$, neglecting only terms of higher order than ρ.

We consider for this purpose the function

$$k(z;x) = g^*(z^*(z);x^*(x)) - g(z;x) \tag{10}$$

with z and x in $D - \sum_{i=0}^{m} \tilde{k}_i$, and with $z^*(z)$ and $x^*(x)$ obtained from them by means of (6). We suppose further that all z_i are situated in D; otherwise, the z_j exterior to D are to be omitted in the summations of all z_i. $k(z;x)$ is harmonic in $D - \sum_{i=0}^{m} \tilde{k}_i$ and is a uniform function of z; for $V_{\rho q}$ transforms this domain into $D^* = \sum_{i=0}^{m} \tilde{k}_i$ and there $g^*(z^*;x^*)$ is defined and harmonic. The pole at $z = x$ is cancelled in the difference $g^* - g$. Therefore by Green's identity

$$k(z;x) = \frac{1}{2\pi}\int_{\mathrm{P}} \{k(t;x)\,\frac{\partial}{\partial n}\,g(z;t) - g(z;t)\,\frac{\partial}{\partial n}\,k(t;x)\}ds, \tag{11}$$

P = boundary of $D - \sum_{i=0}^{m} \tilde{k}_i$, the normal n pointing into the interior of the considered domain.

On the boundary of D, the integrand vanishes, since $k(t;x)$ and $g(z;t)$ do so. Therefore, we have to calculate the integral along the circles k_i only. We introduce (10) into (11) and take into account the identity

$$\frac{1}{2\pi}\int_{k_i} \{g(z;t)\,\frac{\partial}{\partial n}\,g(t;x) - g(t;x)\,\frac{\partial}{\partial n}\,g(z;t)\}ds = 0 \tag{12}$$

Identity (12) results from the fact that $g(z;t)$ and $g(t;t)$ are harmonic in $\tilde{k}_i$ in view of the general identity

$$\int_{k_i} \{u(t)\,\frac{\partial}{\partial n}\,v(t) - v(t)\,\frac{\partial}{\partial n}\,u(t)\}ds = 0, \tag{13}$$

which is valid for each couple of functions $u(z)$ and $v(z)$ which are harmonic in $\tilde{k}_i$. There remains, therefore, only the calculation of

$$S_i = \frac{1}{2\pi}\int_{k_i} \{g^*(t^*(t);x^*)\,\frac{\partial}{\partial n}\,g(z;t) - g(z;t)\,\frac{\partial}{\partial n}\,g^*(t^*(t);x^*)\}ds. \tag{14}$$

We put $t = z_i + re^{i\phi}$; then

$$(14') \qquad S_i = \frac{1}{2\pi}\int_0^{2\pi} \{g^*[t + \frac{a_i\rho}{re^{i\phi}} + \rho y(re^{i\phi}); x^*] \frac{\partial}{\partial r}\ g(z;t)$$

$$- g(z;t) \frac{\partial}{\partial r}\ g^*[t + \frac{a_i\rho}{re^{i\phi}} + py(re^{i\phi}); x^*] \cdot rd\phi$$

where $y(re^{i\phi})$ denotes an analytic function of its argument $re^{i\phi}$. Let $p^*(u;v)$ and $p(u;v)$ be such analytic functions that $\mathcal{R}\{p^*(u;v)\} = g^*(u;v)$ and $\mathcal{R}\{p(u;v)\} = g(u;v)$. Denoting by $p^{*\prime}$, p' the derivatives of the corresponding functions with respect to the first argument, we get by development into series

$$(15) \quad S_i = \frac{1}{2\pi} \int_0^{2\pi} \{g^*(t;x^*) + \mathcal{R}[p^{*\prime}(t;x^*)(\frac{a_i\rho}{re^{i\phi}} + \rho y(re^{i\phi}))]$$

$$+ O(\rho^2)\} \frac{\partial}{\partial r}\ g(z;t) \cdot rd\phi$$

$$- \frac{1}{2\pi} \int_0^{2\pi} g(z;t)\{\frac{\partial}{\partial r}\ (g^*(t;x^*) + \mathcal{R}[p^{*\prime}(t;x^*)(\frac{a_i\rho}{re^{i\phi}} + py(re^{i\phi})])$$

$$+ O(\rho^2)\} \cdot rd\phi$$

with $|O(\rho^2)| < C|\rho|^2$. (15) can be simplified by means of the identity (13); we apply it with $v(t) = g(z;t)$, $u(t) = g^*(t;x^*)$, then with $v(t) = g(z;t)$, $u(t) = \mathcal{R}\{p^*(t;x^*)y(t - z_i)\}$ and at last with $v(t) = g(z;t)$, $u(t) = \mathcal{R}\{[p^{*\prime}(t;x^*) - p^{*\prime}(z_i;x^*)]\frac{a_i\rho}{re^{i\phi}}\}$.

Then

$$(16) \quad S_i = \frac{1}{2\pi} \int_0^{2\pi} \mathcal{R}\{p^{*\prime}(z_i;x^*) \frac{a_i\rho}{e^{i\phi}}\} \frac{\partial}{\partial r}\ g(z;t)d\phi$$

$$+ \frac{1}{2\pi} \int_0^{2\pi} g(z;t)\,\mathcal{R}\{p^{*\prime}(z_i;x^*)\frac{a_i\rho}{re^{i\phi}}\}d\phi + O(\rho^2).$$

Now, $g(z;t) = g(t;z)$ can be developed into a series of powers of $re^{i\phi}$; thus

$$(17) \quad g(t;z) = g(z_i + re_i^{i\phi}z) = g(z_i;z) + \mathcal{R}\{\sum_{\nu=1}^{\infty} \frac{1}{\nu!}\ p^{(\nu)}(z_i;z)r^\nu e^{i\nu\phi}\},$$

if r is sufficiently small. Therefore

$$(18) \quad S_i = \frac{1}{2\pi}\int_0^{2\pi} \mathcal{R}\{p^{*\prime}(z_i;x^*)\ \frac{a_i\rho}{e^{i\phi}}\}\mathcal{R}\{\sum_{\nu=1}^{\infty}\frac{1}{(\nu-1)!}\ p^{(\nu)}(z_i;z)r^{\nu-1}e^{i\nu\phi}\}d\phi$$

$$+ \frac{1}{2\pi}\int_0^{2\pi} [g(z_i;z) + \mathcal{R}\{\sum_{\nu=1}^{\infty} \frac{1}{\nu!}\ p^{(\nu)}(z_i;z)r^\nu e^{i\nu\phi}\}]$$

$$\times \mathcal{R}\{p^{*\prime}(z_i;x^*)\ \frac{a_i\rho}{re^{i\phi}}\}d\phi + O(\rho^2).$$

Putting $\rho a_i p^{*\prime}(z_i; x^*) = Ae^{i\tau}$ and $p'(z_i; z) = Be^{i\sigma}$, and using the formulas expressing the orthogonality of the trigonometric functions, we get

$$(19)\quad S_i = \frac{1}{2\pi}\int_0^{2\pi} 2AB\cos(\tau-\phi)\cos(\sigma+\phi)d\phi + O(\rho^2)$$
$$= \frac{AB}{2\pi}\int_0^{2\pi}[\cos(\sigma+\tau)+\cos(\tau-\sigma-2\phi)]d\phi + O(\rho^2)$$
$$= AB\cos(\sigma+\tau) + O(\rho^2) = \Re\{\rho a_i p^{*\prime}(z_i; x^*)p'(z_i; z)\} + O(\rho^2).$$

Introduce these values into (11) and note that because of the differentiation in the direction of the inner normal n each integral S_i is to be taken positive. We get

$$(20)\quad g^*(z^*(z); x^*(x)) = g(z; x) + \Re\{\rho\sum_{i=0}^{m} a_i p^{*\prime}(z_i; x^*)\ p'(z_i; z)\} + O(\rho^2).$$

Putting $z^* = z + \rho q(z)$, $x^* = x + \rho q(x)$, we develop $g^*(z, x)$ into a series,

$$(21)\quad g^*(z; x) = g(z; x) + \Re\{\rho(\sum_{i=0}^{m} a_i p^{*\prime}(z_i; x^*)p'(z_i; z)$$
$$- q(z)p^{*\prime}(z; x^*) - q(x)p^{*\prime}(x; z^*))\} + O(\rho^2).$$

It is then obvious that $p^{*\prime}(u; v) = p'(u; v) + O(\rho)$, and therefore we get the final equality

$$(22)\quad g^*(z; x) = g(z; x) + \Re\{\sum_{i=0}^{m} a_i p'(z_i; x)p'(z_i; z)$$
$$- q(z)p'(z; x) - q(x)p'(x; z))\} + O(\rho^2).$$

Thus, we have indeed an expression for $g^*(z; x)$ in terms of $g(z; x)$ with the desired approximation.

It is to be pointed out once more that (22) holds only if all points z_i are situated in D. If some z_i are not in D, the summation is to be applied to the $z_i \subset D$ only.

(22) remains valid, even if D has arbitrary boundaries, if there exists at least one boundary continuum. The variation $V_{\rho q}$ transforms D into an analogous domain D^* and both functions $g(z; x)$ and $g^*(z; x)$ exist. Further, there is easily found a sequence of domains D_n on $\mathfrak{R}$, converging to its kernel D in the Carathéodory sense, and whose boundary is composed of analytic curves. By $V_{\rho q}$ we attach to them domains D^*_n of the same type, which converge to D^*. For the corresponding Green's functions $g_n(z; x) \to g(z; x)$ and $g^*_n(z; x) \to g^*(z; x)$ hold uniformly in each inner partial domain of D and D^*. The functions g^*_n and g_n are connected by formula (22), the residual

member $O(\rho^2)$ of which can be evaluated uniformly for each n, since it contains only $g_n(z;x)$, $g^*_n(z;x)$ and their derivatives, all of which converge to the corresponding values of $g(z;x)$, $g^*(z;x)$ etc. (z and x being situated in the interior of D and D^*). Hence, we get in the limit as $n \to \infty$ the formula (22) for $g^*(z;x)$ and $g(z;x)$.

4. The variation of the representing function belonging to D. We suppose now that D is a simply connected domain on $\mathfrak{R}$ so that there exists, therefore, a function $z = f(\zeta)$, which is analytic in $|\zeta| < 1$ and maps this domain in a one-to-one manner on D. Let $z = 0$ be an interior point of D; then $f(\zeta)$ is fixed unambiguously by the conditions $f(0) = 0, f'(0) > 0$. $f(\zeta)$ will be called the representing function belonging to D.

We submit D to a variation $V_{\rho q}$ and get a new domain D^*; by means of (22) we can express the representing function $f^*(\zeta)$ belonging to D^* in terms of $f(\zeta)$, neglecting only terms of a higher order than ρ.

The inverse functions $\phi(z)$ and $\phi^*(z)$ of $f(\zeta)$ and $f^*(\zeta)$ are analytic and univalent in D and D^*, respectively, and map these domains on $|\zeta| < 1$. We have further $\phi(0) = \phi^*(0) = 0$. Hence

$$(23) \qquad g(z;0) = -\log|\phi(z)|; \qquad g^*(z;0) = -\log|\phi^*(z)|.$$

Indeed, the right hand side terms are harmonic in D and D^*, respectively, except at the point $z = 0$ where they become infinite as $-\log|z|$; and they vanish on the boundary of their corresponding domains.

In view of $p(z;0) = -\log\phi(z)$, and because of

$$(23') \qquad p(z;t) = -\log\frac{\phi(z)-\phi(t)}{1-\overline{\phi(t)}\phi(z)},$$

we get from (22) and (23)

$$(24) \quad \log|\phi^*(z)| = \log|\phi(z)| - \mathcal{R}\left\{\rho\Big(\sum_{i=0}^{m} a_i \frac{\phi'(z_i)}{\phi(z_i)} \times \left[\frac{\phi'(z_i)}{\phi(z_i)-\phi(z)} + \frac{\overline{\phi(z)}\phi'(z_i)}{1-\overline{\phi(z)}\phi(z_i)}\right] + q(z)\frac{\phi'(z)}{\phi(z)} - q(0) \times \left[\frac{\phi'(0)}{\phi(0)-\phi(z)} + \frac{\overline{\phi(z)}\phi'(0)}{1-\overline{\phi(z)}\phi(0)}\right]\Big)\right\} + O(\rho^2).$$

Completing both sides of this equation to analytic functions, we get

$$(25) \quad \log\phi^*(z) = \log\phi(z) - \rho\Big(\sum_{i=0}^{m} a_i \frac{\phi'(z_i)^2}{\phi(z_i)[\phi(z_i)-\phi(z)]} + q(z)\frac{\phi'(z)}{\phi(z)} + q(0)\frac{\phi'(0)}{\phi(z)}\Big) - \bar{\rho}\Big(\sum_{i=0}^{m} \bar{a}_i \frac{\overline{\phi'(z_i)}^2\phi(z)}{\overline{\phi(z_i)}[1-\overline{\phi(z_i)}\phi(z)]} - \overline{q(0)\phi'(0)}\phi(z)\Big) + i|\rho|C + O(\rho^2),$$

where C denotes a real constant determined by the condition $\phi'(0) > 0$, $\phi^{*\prime}(0) > 0$. We may bring (25) to the form

$$(25') \quad \phi^*(z) = \phi(z) - \rho\Big(\sum_{i=0}^{m} a_i \frac{\phi'(z_i)^2\phi(z)}{\phi(z_i)[\phi(z_i) - \phi(z)]} + q(0)\phi'(0) + q(z)\phi'(z)\Big)$$

$$- \bar{\rho}\Big(\sum_{i=0}^{m} \bar{a}_i \frac{\overline{\phi'(z_i)}^2\phi(z)^2}{\overline{\phi(z_i)}[1 - \overline{\phi(z_i)}\phi(z)]} - \overline{q(0)\phi'(0)}\phi(z)^2\Big) + i\,|\,\rho\,|\,C\phi(z) + O(\rho^2)$$

and put $\phi(z) = \zeta$, $\phi'(z) = f'(\zeta)^{-1}$, $\phi(z_i) = \zeta_i$. We have, further,

$$(26) \quad f(\zeta) = z = f^*(\phi^*(z)) = f^*(\phi(z)) + f^{*\prime}(\phi(z))\,[-\rho(\cdots)$$
$$- \bar{\rho}(\cdots) + i\,|\,\rho\,|\,C\phi(z)] + O(\rho^2),$$

and since $f^{*\prime}(\zeta) = f'(\zeta) + O(\rho)$, we obtain finally

$$(27) \quad f^*(\zeta) = f(\zeta) + \rho f'(\zeta)\Big(\sum_{i=0}^{m} a_i \frac{\zeta}{\zeta_i f'(\zeta_i)^2(\zeta_i - \zeta)} + \frac{q(0)}{f'(0)} + \frac{q(f(\zeta))}{f'(\zeta)}\Big)$$
$$+ \bar{\rho} f'(\zeta)\Big(\sum_{i=0}^{m} \bar{a}_i \frac{\zeta^2}{\bar{\zeta}_i \overline{f'(\zeta_i)}^2(1 - \bar{\zeta}_i\zeta)} - \frac{\overline{q(0)}}{\overline{f'(0)}}\zeta^2\Big)$$
$$- i\,|\,\rho\,|\,C\zeta f'(\zeta) + O(\rho^2).$$

Thus our aim is attained; for the sake of completeness, we compute C explicitly. We have

$$(27') \quad f^{*\prime}(0) = f'(0)\Big[1 + \rho\sum_{i=0}^{m} \frac{a_i}{\zeta_i^2 f'(\zeta_i)^2} + \rho q'(0) - i\,|\,\rho\,|\,C\Big] + O(\rho^2).$$

$f^{*\prime}(0)$ being positive, the multiplier of $f'(0)$ must be positive. Hence we get

$$(27'') \quad C = \mathfrak{J}\Big\{\frac{\rho}{|\,\rho\,|}\Big(\sum_{i=0}^{m} \frac{a_i}{\zeta_i^2 f'(\zeta_i)^2} + q'(0)\Big)\Big\}.$$

5. An application to the coefficient problem for univalent functions. A function $z = f(\zeta)$ is called univalent in $|\,\zeta\,| < 1$, if it maps this domain in a one-to-one manner on a domain D of the z-plane. The z-plane plays here the part of the closed R. S. $\mathfrak{R}$, containing D. According to the discussion of **2** each variation $V_{\rho q}$ transforms the z-plane into itself, since there is only one closed one-sheeted R. S. over the z-plane. For our purposes it is sufficient to choose

$$(28) \quad q(z) = \frac{z}{z - z_0}$$

with $z_0 \subset D$.

We seek the functions with the normalization (1) which are regular and

univalent in $|\zeta| < 1$, and which possess an n-th coefficient a_n with the largest possible absolute value. Let $f_n(\zeta)$ be such a function and let D_n be the domain, on which it maps $|\zeta| < 1$. The variation $V_{\rho q}$ with $q(z)$ defined as in (28) transforms D_n into a new domain D^*_n with the representing function

$$(29)\quad f^*_n(\zeta) = f_n(\zeta) + \rho\{f'_n(\zeta)\,\frac{f_n(\zeta_0)}{\zeta_0^2 f'_n(\zeta_0)^2}\,\frac{\zeta}{1-\zeta/\zeta_0} + \frac{f_n(\zeta)}{f_n(\zeta)-f_n(\zeta_0)}\}$$
$$+ \bar{\rho} f'_n(\zeta)\,\frac{\overline{f_n(\zeta_0)}}{\bar{\zeta}_0 \overline{f'_n(\zeta_0)^2}}\,\frac{\zeta^2}{1-\bar{\zeta}_0\zeta} - i\,|\rho|\,C\zeta f'_n(\zeta) + O(\rho^2).$$

$f^*_n(\zeta)$ is also regular and univalent in $|\zeta| < 1$, but it has not the normalization (1). This holds, however, for $f^*_n(\zeta) f^{*\prime}_n(0)^{-1}$, and in view of the extremal property of $f_n(\zeta)$ we obtain

$$(30)\quad |a^*_n f^{*\prime}_n(0)^{-1}| = |a_n + \rho\,\frac{f_n(\zeta_0)}{\zeta_0^2 f'_n(\zeta_0)^2}\,((n-1)a_n + (n-1)a_{n-1}\cdot\frac{1}{\zeta_0}$$
$$+ (n-2)a_{n-2}\,\frac{1}{\zeta_0^2} + \cdots + \frac{1}{\zeta_0^{n-1}}) - \rho\,\frac{1}{f_n(\zeta_0)}\,P_n\left(\frac{1}{f_n(\zeta_0)}\right)$$
$$+ \bar{\rho}\,\frac{f_n(\zeta_0)}{\overline{\zeta_0^2 f'_n(\zeta_0)^2}}\,((n-1)a_{n-1}\bar{\zeta}_0 + (n-2)a_{n-2}\bar{\zeta}_0^2 + \cdots + \bar{\zeta}_0^{n-1})$$
$$- i\,|\rho|\,C(n-1)a_n + O(\rho^2)| \leq |a_n|$$

where $P_n(x)$ is defined by (2). Without loss of generality we may suppose $a_n > 0$; then (30) implies

$$(30')\quad \Re\left\{\rho\left[\frac{f_n(\zeta_0)}{\zeta_0^2 f'_n(\zeta_0)^2}\left((n-1)a_n + (n-1)\frac{a_{n-1}}{\zeta_0} + \cdots + \frac{1}{\zeta_0^{n-1}}\right)\right.\right.$$
$$\left.\left. - \frac{1}{f_n(\zeta_0)}\,P_n\left(\frac{1}{f_n(\zeta_0)}\right)\right] + \bar{\rho}\,\frac{\overline{f_n(\zeta_0)}}{\overline{\zeta_0^2 f'_n(\zeta_0)^2}}\left((n-1)a_{n-1}\bar{\zeta}_0 + \cdots \bar{\zeta}_0^{n-1}\right)\right\}$$
$$+ O(\rho^2) \leq 0.$$

This holds for each value $\rho = |\rho|\,e^{i\psi}$. If we denote the coefficients of ρ and $\bar{\rho}$ by M and N respectively, and pass to the limit $|\rho| = 0$ in (30'), we find

$$(30'')\qquad \Re\{e^{i\psi}M + e^{-i\psi}N\} \leq 0 \quad \text{for} \quad 0 \leq \psi \leq 2\pi.$$

Since obviously $\Re\{e^{i\psi}M - e^{-i\psi}\bar{M}\} = 0$, (30'') yields

$$(30''')\qquad \Re\{e^{-i\psi}(\bar{M} + N)\} \leq 0$$

for each value of ψ. This is only possible in the case

$$(30^{\mathrm{IV}})\qquad \bar{M} + N = 0.$$

Introducing the values of M and N into (30^{IV}), we get the equality

$$(31)\qquad \frac{f_n(\zeta_0)}{\zeta_0^2 f'_n(\zeta_0)^2}\left(\frac{1}{\zeta_0^{n-1}}+\frac{2a_2}{\zeta_0^{n-2}}+\cdots+\frac{(n-1)a_{n-1}}{\zeta_0}\right.$$

$$\left.+(n-1)a_n+(n-1)\bar{a}_{n-1}\zeta_0+\cdots\zeta_0^{n-1}\right)=\frac{1}{f_n(\zeta_0)}\;P_n\left(\frac{1}{f_n(\zeta_0)}\right),$$

i. e.

$$(31')\qquad \frac{\zeta_0^2 f'_n(\zeta_0)^2}{f_n(\zeta_0)^2}\;P_n\left(\frac{1}{f_n(\zeta_0)}\right)=\frac{1}{\zeta_0^{n-1}}+\frac{2a_2}{\zeta_0^{n-2}}+\cdots$$

$$+(n-1)a_n+\cdots+2\bar{a}_2\zeta_0^{n-2}+\zeta_0^{n-1}.$$

(31′) holds for each value ζ_0 in $|\zeta|<1$; hence $f_n(\zeta)$ must satisfy this differential equation everywhere and can be continued with its aid over its whole domain of existence. In particular (31′) implies that D_n is bounded by analytic curves satisfying the differential equation (3). For let $z=f(e^{i\tau})$ be the parametric representation of the boundary curve; then (31′) yields

$$(31'')\qquad \frac{z'^2}{z^2}\;P_n\left(\frac{1}{z}\right)\leq 0$$

which is equivalent to (3) in view of the supposition $a_n>0$.

Our method shows also that there are no points in the z-plane which are exterior to D_n. Suppose, for example, that z_0 were an exterior point of D_n, then use it for the function (28). We transform D_n into D^*_n by $V_{\rho q}$; but (20) has now the form $g^*(z^*;x^*)=g(z^*;x)$, since there is no z_i in D_n. An easy computation shows that the representing function belonging to D^*_n is

$$(29')\qquad f^*_n(\zeta)=f_n(\zeta)+\rho\,\frac{f_n(\zeta)}{f_n(\zeta)-z_0}-i\,|\,\rho\,|\,C\zeta f'_n(\zeta)+O(\rho^2).$$

Therefore we have for each sufficiently small ρ

$$(32)\qquad |\,a^*_n f^*_n(0)^{-1}\,|=|\,a_n-\rho z_0 P_n\left(\frac{1}{z_0}\right)-i\,|\,\rho\,|\,C(n-1)a_n+O(\rho^2)\,|\leq a_n$$

and in the same way as above we deduce that $\frac{1}{z_0}\;P_n\left(\frac{1}{z_0}\right)=0$. This equation has a finite number of roots z_0. But were there one point z_0 exterior to D_n, there would exist an infinity of such points. This is impossible, and so we have proved that D_n covers all the plane, analytic slits excepted.

The above derivation of the differential equation (3) gives us the explicit formula (31′) for $\frac{\zeta^2 f'_n(\zeta)^2}{f_n(\zeta)^2}\;P_n\left(\frac{1}{f_n(\zeta)}\right)$. If (3) is proved in an alternative

way, Schwarz' principle of reflection implies that the above expression must be a rational function of ζ. By (31′) we expressed the coefficients of this rational function in a simple form by those of $f_n(\zeta)$; in the following paragraph we shall prove this representation in an alternative way.

6. The relation between $P_n(x)$ and the Faber polynomials. We denote by $H_n(\zeta; f)$ the expression $\frac{\zeta^2 f'(\zeta)^2}{f(\zeta)^2} P_n\left(\frac{1}{f(\zeta)}\right)$. If $z = f_n(\zeta)$ is the function with maximal $a_n > 0$ considered in the last section, we see from (3) that $H_n(\zeta; f_n(\zeta))$ tends to real values, if ζ approaches the unit circle. Hence the principle of reflection yields

$$H_n((1/\zeta); f_n) = \overline{H(\zeta; f_n)}. \tag{33}$$

This means that $H_n(\zeta; f)$ necessarily has the form $\frac{1}{\zeta^{n-1}} + \frac{\alpha_2}{\zeta^{n-2}} + \cdots + \frac{\alpha_{n-1}}{\zeta} + a_n + \bar{a}_{n-1}\zeta + \cdots + \zeta^{n-2}$. Hence it suffices to compute the first n coefficients of H_n. Now it can be shown that, if the power series $f(\zeta) = \zeta + a_2\zeta^2 + \cdots + a_n\zeta^n + \cdots$ converges near $\zeta = 0$, the corresponding $H_n(\zeta; f)$ has there the development

$$H_n(\zeta; f) = \frac{1}{\zeta^{n-1}} + \frac{2a_2}{\zeta^{n-2}} + \frac{3a_3}{\zeta^{n-3}} + \cdots + \frac{(n-1)a_{n-1}}{\zeta} + (n-1)a_n + \zeta(\cdots). \tag{34}$$

This identity will be deduced from another, which connects the $P_n(x)$ with the Faber polynomials, whose important rôle in the theory of univalent functions is well known.[2]

If $f(\zeta)$ is a power series of the form (1) converging in a neighborhood of $\zeta = 0$, $F_n(x)$ is called the n-th Faber polynomial belonging to $f(\zeta)$, if it is a polynomial of the n-th degree such that

$$F_n\left(\frac{1}{f(\zeta)}\right) = \frac{1}{\zeta^n} + A_n\zeta + B_n\zeta^2 + \cdots. \tag{35}$$

For each $f(\zeta)$ and for each n there exists exactly one $F_n(x)$. All the polynomials $F_n(x)$ belonging to a fixed function $f(\zeta)$ can be represented by means of a generating function. To show this, consider the function

$$L(\zeta; \eta) = \log(1 - f(\zeta)/f(\eta)) - \log(1 - \zeta/\eta) - \log\frac{f(\zeta)}{\zeta}, \tag{36}$$

[2] H. Grunsky, "Koeffizientenbedingungen für schlicht abbildende meromorphe Funktionen," *Math. Zeits.*, vol. 45 (1939), pp. 29-61.

which in the neighborhood of $\zeta = 0$, $\eta = 0$ is an analytic function of its two arguments. In a sufficiently small circle around $\zeta = 0$, $f(\zeta)$ is univalent; thus $(f(\zeta) - f(\eta))(\zeta - \eta)^{-1}$ is regular and non-zero in this circle and $L(\zeta;\eta)$ is regular in it. Therefore $L(\zeta;\eta)$ can be developed in a series of powers of ζ and η

$$L(\zeta;\eta) = \sum_{\nu,\mu=1}^{\infty} c_{\nu\mu}\zeta^{\nu}\eta^{\mu}. \tag{37}$$

If, on the other hand, we develop $L(\zeta;\eta)$ in powers of ζ for $\eta \neq 0$ fixed, we get

$$L(\zeta;\eta) = -\sum_{\nu=1}^{\infty} \frac{1}{\nu} F_{\nu}\left(\frac{1}{f(\eta)}\right)\zeta^{\nu} + \sum_{\nu=1}^{\infty} \frac{1}{\nu} \frac{\zeta^{\nu}}{\eta^{\nu}}. \tag{37'}$$

Here $F_{\nu}(x)$ is a polynomial of the ν-th degree, and a comparison of (37) and (37′) yields

$$F_n\left(\frac{1}{f(\eta)}\right) = \frac{1}{\eta^n} - n\sum_{\mu=1}^{\infty} c_{n\mu}\zeta^{\mu}. \tag{37''}$$

Hence, $F_n(x)$ is indeed the n-th Faber polynomial belonging to $f(\zeta)$, and all $F_n(x)$ are obtained by means of the generating function

$$\log(1 - xf(\zeta)) - \log\frac{f(\zeta)}{\zeta} = -\sum_{\nu=1}^{\infty} \frac{1}{\nu} F_{\nu}(x)\zeta^{\nu}. \tag{38}$$

From (38) we get by differentiating with respect to x

$$\frac{f(\zeta)}{1 - xf(\zeta)} = \sum_{\nu=1}^{\infty} \frac{1}{\nu} F'_{\nu}(x)\zeta^{\nu} \tag{39}$$

and hence by comparison with (2) [3]

$$\frac{1}{n} F'_n(x) = a_n + P_n(x). \tag{40}$$

The identity (34) is now easily obtained. Differentiation of (35) with respect to ζ and the application of (40) yield

$$\frac{f'(\zeta)}{f(\zeta)^2}\left[a_n + P_n\left(\frac{1}{f(\zeta)}\right)\right] = \frac{1}{\zeta^{n+1}} - \frac{A_n}{n} + \cdots \tag{41}$$

and hence

$$\frac{f'(\zeta)^2}{f(\zeta)^2} P_n\left(\frac{1}{f(\zeta)}\right) = \frac{1}{\zeta^{n+1}} + \frac{2a_2}{\zeta^n} + \cdots + \frac{(n-1)a_{n-1}}{\zeta^3} + \frac{(n-1)a_n}{\zeta^2} + \frac{(n+1)a_{n+1} - 2a_2a_n}{\zeta} + \cdots. \tag{41'}$$

From this the desired identity follows immediately.

---

[3] The author found this relation after a most helpful discussion with the late Prof. Schur, whose investigations on Faber polynomials will be published shortly.

In the case of the extremal function $f_n(\zeta)$, we know from (31′) that the coefficient of ζ in the development of $H_n(\zeta; f_n)$ is equal to $(n-1)\bar{a}_{n-1}$; on the other hand, (41′) gives for this coefficient the value $(n+1)a_{n+1} - 2a_2a_n$. Therefore, in this case

$$(n+1)a_{n+1} = 2a_2a_n + (n-1)\bar{a}_{n-1}, \tag{42}$$

a relation established by Marty.

7. A coefficient problem from the theory of *p*-valued functions. We shall now show, by means of a special problem, how the above method of variation is to be applied in the theory of p-valued functions. Let

$$f(\zeta) = \zeta + a_2\zeta^2 + a_3\zeta^3 + \cdots + a_n\zeta^n + \cdots \tag{1}$$

be regular in $|\zeta| < 1$, taking on the value zero only at $\zeta = 0$, and taking on each other value at most p times. All these functions $f(\zeta)$ are known to form a compact family, and therefore we may seek those functions $f_n(\zeta)$, for which $|a_n|$ possesses the largest value in the family.

Each function of the family maps $|\zeta| < 1$ on a domain D situated on a p-sheeted R. S. over the z-plane. Let us consider the subclass $\mathfrak{F}_g$ of all functions within the family, the domain D of which can be stretched over a p-sheeted R. S. with genus $\leqq g$. $\mathfrak{F}_g$ also is a compact family, and we first consider the coefficent problem for this. Knowing the maximum $\mu_{n,g}$ for $|a_n|$ with respect to $\mathfrak{F}_g$, we get the maximum μ_n for $|a_n|$ with respect to the whole family by the equality

$$\mu_n = \lim_{g=\infty} \mu_{n,g}. \tag{43}$$

It is sufficient, therefore, to restrict ourselves to the coefficient problem for the family $\mathfrak{F}_g$; our method will enable us to make some assertions concerning the extremal functions.

Let, therefore, g be fixed and let $f_n(\zeta)$ be a function of the family $\mathfrak{F}_g$ with maximal $|a_n|$. This function determines a domain D_n on a R. S. $\mathfrak{R}_n$ which covers the point 0 exactly once, and does not cover the point ∞. $\mathfrak{R}_n$ is p-sheeted and of genus $\gamma \leqq g$. We choose a function $q(z)$ uniform on $\mathfrak{R}_n$, regular and finite in each point with coördinate $z = \infty$, and regular and vanishing in each point with coördinate $z = 0$. We apply the variation $V_{\rho q}$ to D_n. The new domain D^*_n belongs to the representing function $f^*_n(\zeta)$ obtained from $f_n(\zeta)$ by means of (27) and possessing the n-th coefficient

$$(44)\quad a^*_n = a_n + \rho\Big(\sum_{i=0}^{m} \frac{a_i}{\zeta_i^2 f'_n(\zeta_i)^2}\Big[na_n + (n-1)\frac{a_{n-1}}{\zeta_i} + \cdots + \frac{1}{\zeta_i^{n-1}}\Big]$$

$$+ \{q(f_n(\xi))\}_{n\text{-th coeff.}}\Big) + \bar{\rho}\Big(\sum_{i=0}^{m} \frac{\bar{a}_i}{\bar{\zeta}_i^2 \overline{f'_n(\zeta_i)^2}}\Big[(n-1)a_{n-1}\bar{\zeta}_i + \cdots + \bar{\zeta}_i^{n-1}\Big]\Big)$$

$$- i\,|\,\rho\,|\,Cna_n + O(\rho^2).$$

$V_{\rho q}$ preserves the genus and the number of sheets; hence $f^*_n(\zeta) f^*_n(0)^{-1}$ is a function of the family $\mathfrak{F}_g$ and therefore

$$(45)\quad |\,a^*_n f^{*\prime}_n(0)^{-1}\,| = |\,a_n + \rho\Big(\sum_{i=0}^{m} \frac{a_i}{\zeta_i f'_n(\zeta_i)^2}\Big[(n-1)a_n + (n-1)\frac{a_{n-1}}{\zeta_i}$$

$$+ \cdots + \frac{1}{\zeta_i^{n-1}}\Big] + \{q(f_n(\xi))\}_{n\text{-th coeff.}} - a_n q'(f_n(0))$$

$$+ \bar{\rho}\sum_{i=0}^{m} \frac{\bar{a}_i}{\bar{\zeta}_i^2 \overline{f'_n(\zeta_i)^2}}\Big[(n-1)a_{n-1}\bar{\zeta}_i + \cdots + \bar{\zeta}_i^{n-1}\Big] - i\,|\,\rho\,|\,C(n-1)a_n$$

$$+ O(\rho^2)\,| \leq |\,a_n\,|\,.$$

Supposing, once more, a_n to be positive, we get finally by the conclusions of **5** the equation

$$(46)\quad \sum_{i=0}^{m} a_i \frac{1}{\zeta_i^2 f'_n(\zeta_i)^2}\Big(\frac{1}{\zeta_i^{n-1}} + \frac{2a_2}{\zeta_i^{n-2}} + \cdots + \frac{(n-1)a_{n-1}}{\zeta_i} + (n-1)a_n$$

$$+ (n-1)\bar{a}_{n-1}\zeta_i + \cdots + \zeta_i^{n-1}\Big)$$

$$= a_n q'(f_n(0)) - \{q(f_n(\xi))\}_{n\text{-th coeff.}}\,.$$

There still remains the problem of examining all the functions $q(z)$ of the type considered, which are uniform on $\mathfrak{R}_n$. We get such a function in the easiest way by putting $q(z) = \dfrac{z}{z - z_0}$ supposing only that no boundary point of D_n and no branch point of R_n has the coördinate z_0. Then (46) yields

$$(47)\quad \sum_{i=0}^{m} \frac{f_n(\zeta_i)}{\zeta_i^2 f'_n(\zeta_i)^2}\,(1/\zeta_i^{n-1} + \cdots + (n-1)a_n + \cdots + \zeta_i^{n-1})$$

$$= \frac{1}{z_0} P_n\Big(\frac{1}{z_0}\Big), \quad f_n(\zeta_i) = z_0.$$

This equality, however, is not the most general assertion concerning $f_n(\zeta)$. To get such an assertion we shall use in the following section a more general type of functions $q(z)$ uniform on $\mathfrak{R}_n$.

8. The variation function $q(z)$, uniform on $\mathfrak{R}_n$. Let $\mathfrak{R}_n$ be the R. S. over which the domain D_n, belonging to the extremal function $f_n(\zeta)$, is extended. γ denoting the genus of $\mathfrak{R}_n$, there exist 2γ loop-cuts $A_1, B_1, \cdots, A_\gamma, B_\gamma$ on the surface, with the following properties: Each loop-cut A_i meets the cut B_i in exactly one point π_i, but meets no other loop-cut; analogously each B_i meets only the corresponding A_i. Choose a fixed point P on $\mathfrak{R}_n$ and connect it with each π_i by a Jordan curve C_i, which has no points in common with the loop-cuts and the other C_k, the points π_i and P excepted. The A_i, B_i, C_i $(i = 1, 2, \cdots \gamma)$ yield a canonical resolution of $\mathfrak{R}_n$.

A function $w(z)$ is called an integral of the first kind on $\mathfrak{R}_n$, if it possesses a finite and uniform derivative with respect to the local parameter at each point of the R. S. $w(z)$ itself is of course not uniform on $\mathfrak{R}_n$, if it is not a constant, but it increases by certain constant numbers, the periods, if its argument crosses the cuts A_i, B_i. There exist exactly $\gamma + 1$ integrals of the first kind 1, $w_1(z), \cdots, w_\gamma(z)$ on $\mathfrak{R}_n$, which are linearly independent. We can suppose them normalized in such a manner that $w_i(z)$ has the period δ_{ik} with respect to the crossing of A_k, with $\delta_{ik} = 0$ if $i \neq k$ and $= 1$ if $i = k$. These $w_i(z)$ are called the normal integrals of the first kind (transcendent normalization). $\tau(z; x)$ is called an integral of the second kind with pole x, if its derivative with respect to the local parameter is regular and uniform everywhere on $\mathfrak{R}_n$, except at the point x where it has a double pole. Since if $\tau(z; x)$ is, then $\tau(z; x) - \sum_{\nu=0}^{\gamma} c_\gamma w_\gamma(z)$ is also an integral of the second kind with the same pole, we can normalize $\tau(z; x)$ in such a way that all its periods with respect to the A_i vanish. If x is not a branch point of $\mathfrak{R}_n$ we may choose in its neighborhood z itself as a local parameter. Then $\tau(z; x)$ is to be chosen so that the development

$$\tau(z; x) = \frac{1}{z - x} + a + b(z - x) + \cdots \text{ holds.} \tag{48}$$

All these conditions together fix $\tau(z; x)$ except for an additive constant; $\tau(z; x)$ is called a normal integral of the second kind.

The periods of $\tau(z; x)$ with respect to the A_k are zero by definition; with respect to the B_k, $\tau(z; x)$ has the periods

$$P_k(x) = -2\pi i \frac{dw_k(x)}{dx} = -2\pi i w'_k(x). \tag{49}$$

$\omega(z; x_1, x_2)$ is called an integral of the third kind on $\mathfrak{R}_n$ if its derivative with respect to the local parameter is uniform and finite everywhere, except

for the points x_1 and x_2, where ω' has two simple poles with the sum of residues zero. Hence $\omega(z; x_1, x_2)$ has logarithmic poles at x_1 and x_2. We normalize $\omega(z; x_1, x_2)$ so that its periods with respect to the A_i vanish and so that the residues of its derivative at x_1 and x_2 are -1 and $+1$ respectively. Then $\omega(z; x_1, x_2)$ will be called a normal integral of the third kind.

Between the normal integrals of the second and third kind we have the important relation

$$\frac{d}{dz}\omega(z; x_1, x_2) = \omega'(z; x_1, x_2) = \tau(x_1; z) - \tau(x_2; z). \tag{50}$$

This equation permits certain assertions concerning the dependence of $\tau(z; x)$ upon its pole x. Examine, for example, $\tau(x_1; z) - \tau(x_2; z)$ in the neighborhood of a branch point z_b of $\mathfrak{R}_n$, which is supposed different from x_1 and x_2. If $t = (z - z_b)^\alpha$ is the local parameter, we have by (50) $\tau(x_1; z) - \tau(x_2; z) = \alpha(z - z_b)^{\alpha-1}\frac{d}{dt}\omega(z; x_1, x_2)$. Now, $\frac{d}{dt}\omega(z; x_1, x_2)$ is an analytic function of t, x_1 and x_2 and can, therefore, be differentiated with respect to x_1 an arbitrary number of times. This shows that each derivative of $\tau(x_1; z)$ with respect to x_1 has at z_b an infinity of order $(z - z_b)^{\alpha-1}$ at most. This fact will be of use to us later.

Having recalled the notions from the theory of Riemann surfaces which are necessary for our investigation, we proceed to the construction of a function $q(z)$, uniform and meromorphic on $\mathfrak{R}_n$, vanishing at $z = 0$ and finite at $z = \infty$. For this purpose, we fix γ points $z_1, z_2, \cdots z_\gamma$ such that the determinant

$$|w'_k(z_i)|_{i,k=1,2,\ldots\gamma} \neq 0. \tag{51}$$

This is always possible, the $w_k(z)$ being linearly independent. The z_i can even be chosen in D_n and different from all the branch points of $\mathfrak{R}_n$. z_0 being also a point of this type, we form

$$q_1(z) = \sum_{\nu=0}^{\gamma} a^*_\nu \tau(z; z_\nu) \tag{52}$$

and determine the a^*_ν in such a way that $q_1(z)$ is uniform on $\mathfrak{R}_n$. In view of (49), this requirement is equivalent to the equations

$$\sum_{\nu=0}^{\gamma} a^*_\nu w'_k(z_\nu) = 0 \qquad k = 1, 2, \cdots \gamma, \tag{53}$$

which have, because of (51), one solution, not taking into account an arbitrary common factor. $q_1(z)$ is uniform on $\mathfrak{R}_n$ and regular everywhere, the simple

poles z_ν excepted. In the neighborhood of z_ν, $q_1(z)$ has the development (5). We form now

$$q(z) = \frac{z}{z-a}\, q_1(z), \tag{54}$$

a being a point in the z-plane which is not a branch point of $\mathfrak{R}_n$ but over which are situated p points of D_n. There exists such a point a, $f_n(\zeta)$ being supposed p-valued; denote by $z_{\gamma+1}, \cdots, z_{\gamma+p}$ the p points of D_n with the coördinate a.

The function $q(z)$ satisfies all our requirements and can be used as the variation function. Thus we get from (46)

$$\sum_{i=0}^{\gamma+p} a_i \chi(\zeta_i) = a_n q'(f_n(0)) - \{q(f_n(\xi))\}_{n\text{-th coeff.}}\ ; \tag{46'}$$

if we put here

$$\chi(\zeta) = \frac{1}{\zeta^2 f'_n(\zeta)^2}\left(\frac{1}{\zeta^{n-1}} + \frac{2a_2}{\zeta^{n-2}} + \cdots + \frac{(n-1)a_{n-1}}{\zeta} + (n-1)a_n + (n-1)\bar{a}_{n-1}\zeta + \cdots + \zeta^{n-1}\right), \tag{46''}$$

$f(\zeta_i) = z_i$ and

$$a_i = \frac{z_i}{z_i - a} \text{ for } i = 0, \cdots \gamma; \quad a_i = a q_1(z_i) \text{ for } i = \gamma+1, \cdots, \gamma+p. \tag{55}$$

By means of (52) we get from (46')

$$\sum_{\nu=0}^{\gamma} a^*{}_\nu \left[\frac{z_\nu}{z_\nu - a}\chi(\zeta_\nu) + a\sum_{i=\gamma+1}^{\gamma+p}\chi(\zeta_i)\tau(z_i; z_\nu) + \left\{\frac{f_n(\xi)}{f_n(\xi)-a}\tau(f_n(\xi); z_\nu)\right\}_{n\text{-th coeff.}} + \frac{a_n}{a}\tau(f_n(0); z_\nu)\right] = 0. \tag{56}$$

The γ sets of $\gamma+1$ numbers $w'_k(z_0), \cdots w'_k(z_\gamma)$ $(k = 1, \cdots \gamma)$ being linearly independent and satisfying the equation (53), we deduce from (56)

$$\frac{z_\nu}{z_\nu - a}\chi(\zeta_\nu) + a\sum_{i=\gamma+1}^{\gamma+p}\chi(\zeta_i)\tau(z_i; z_\nu) + \left\{\frac{f_n(\xi)}{f_n(\xi)-a}\tau(f_n(\xi); z_\nu)\right\}_{n\text{-th coeff.}} + \frac{a_n}{a}\tau(f_n(0); z_\nu) = \sum_{k=1}^{\gamma}\lambda_k w'_k(z_\nu). \tag{57}$$

Since ζ_0 and $z_0 = f_n(\zeta_0)$ are arbitrary, the condition $z_0 \subset D_n$ and $z_0 \neq$ branch point of $\mathfrak{R}_n$ excepted, we can drop the index ν in (57). Applying, further, (47) to $a = z_0$ and $\zeta_{\gamma+1}, \cdots \zeta_{\gamma+p}$, we get

$$(47') \qquad a\sum_{i=\gamma+1}^{\gamma+p}\chi(\zeta_i)+\left\{\frac{f_n(\xi)}{f_n(\xi)-a}\right\}_{n\text{-th coeff.}}+\frac{a_n}{a}=0.$$

Let y be a regular point of $\mathfrak{R}_n$; we multiply (47') by $-\tau(y;z)$ and add the resulting equation to (57). We get

$$(58) \qquad \frac{z}{z-a}\chi(\zeta)+a\sum_{i=\gamma+1}^{\gamma+p}\chi(\zeta_i)(\tau(z_i;z)-\tau(y;z))+$$

$$\left\{\frac{f_n(\xi)}{f_n(\xi)-a}[\tau(f_n(\xi);z)-\tau(y;z)]\right\}_{n\text{-th coeff.}}+\frac{a_n}{a}[\tau(f_n(0);z)-\tau(y;z)]$$

$$=\sum_{k=1}^{\gamma}\lambda_k w'_k(z).$$

Now, we apply equation (50), which shows that $\tau(x;z)-\tau(y;z)$ and all its derivatives with respect to x are analytic functions of z, becoming infinite in the neighborhood of a $(k-1)$-fold branch point z_b no faster than $(z-z_b)^{(1-k)/k}$. Since the n-th coefficient of ξ in the expression $f_n(\xi)(f_n(\xi)-a)^{-1}[\tau(f_n(\xi);z)-\tau(y;z)]$ is a linear combination of $\tau(0;z)-\tau(y;z)$, $\tau'(0;z),\cdots\tau^{(n)}(0;z)$, and since, further, all $w'_k(z)$ behave at z_b like these functions, $\chi(\zeta)$ becomes infinite at ζ_b (corresponding to the $(k-1)$-fold branch point z_b) as $(z-z_b)^{(1-k)/k}$ at most. $\chi(\zeta)$ does not depend on the arbitrary values a and y; hence, in view of (58), it has its only singularities at $z=0$ and at the points ζ_b.

It is easy to find a linear aggregate of uniform normal integrals on $\mathfrak{R}_n$, which behaves at $z=f_n(0)$ as $z\chi(\zeta)$. To do so, we introduce the normal integrals $\tau_\nu(z;f_n(0))$ possessing everywhere on $\mathfrak{R}_n$ a finite and uniform derivative with respect to the local parameter, the point $f_n(0)$ excepted. In the neighborhood of this point,

$$(59) \qquad \tau_\nu(z;f_n(0))=(1/z^\nu)+\alpha+\beta z+\cdots$$

and the periods with respect to the loop-cuts A_i are zero. Further, let

$$(60) \qquad P_n(x)=\sum_{l=2}^{n}a_{nl}x^{l-1}$$

be the polynomial defined in (2). Then, the expression

$$(61) \qquad L_n(f_n(\zeta))=-\sum_{l=2}^{n}\frac{a_{nl}}{l-1}\tau'_{l-1}(f_n(\zeta);f_n(0))$$

has at $\zeta=0$ the same infinity as

$$(61') \quad \sum_{l=2}^{n} a_{nl} f_n(\zeta)^{-l} = f_n(\zeta)^{-1} P_n\left(\frac{1}{f_n(\zeta)}\right)$$
$$= \frac{f_n(\zeta)}{f'_n(\zeta)^2}\left(\frac{1}{\zeta^{n+1}} + \frac{2a^2}{\zeta^n} + \cdots + \frac{(n-1)a_n}{\zeta^2} + \cdots\right)$$

the last equation resulting from (41'). According to (46''), $L_n(f_n(\zeta))$ has at $\zeta = 0$ the same infinity as $z\chi(\zeta)$; at the branch points z_b, $L_n(z)$ becomes infinite as $(z - z_b)^{(1-k)/k}$ at most. Everywhere else $L_n(z)$ is uniform and regular on $\mathfrak{R}_n$. Hence, $z\chi(\zeta) - L_n(z)$ is uniform on $\mathfrak{R}_n$, becomes infinite at the branch points at most as the derivative with respect to z of a function which is regular in the local parameter, and is regular everywhere else with the possible exception of $z = \infty$.

According to (58) $\chi(\zeta)$ is an algebraic function of z which vanishes at $z = \infty$; hence

$$(62) \quad \chi(\zeta) = c_\kappa z^{-\kappa} + c_\lambda z^{-\lambda} + \cdots, \qquad 0 < \kappa < \lambda < \cdots, \quad c_\kappa \neq 0.$$

Let α be a point on the circle $|\zeta| = 1$, for which $f_n(\alpha) = \infty$; we put $z = f_n(\zeta)$ and deduce from (46'') and (62) that

$$(63) \quad C_\mu(\zeta - \alpha)^\mu + C_{\mu+1}(\zeta - \alpha)^{\mu+1} + \cdots$$
$$= \frac{f'_n(\zeta)^2}{f_n(\zeta)^\kappa}(c_\kappa + c_\lambda f_n(\zeta)^{-(\lambda-\kappa)} + \cdots), \quad \mu \geq 0,$$

which means

$$(63') \quad D_\mu(\zeta - \alpha)^{\mu/2} + \cdots = \frac{f'_n(\zeta)}{f_n(\zeta)^{\kappa/2}}\ (d_\kappa + d_\lambda f_n(\zeta)^{-(\lambda-\kappa)} + \cdots).$$

By integration we get, if $\kappa \neq 2$

$$(64) \quad E_\mu(\zeta - \alpha)^{(\mu+2)/2} + \cdots = f_n(\zeta)^{1-(\kappa/2)}(e_\kappa + e_\lambda f_n(\zeta)^{-(\lambda-\kappa)} + \cdots) + \text{Const},$$

while for $\kappa = 2$ we have

$$(64') \quad E_\mu(\zeta - \alpha)^{(\mu+2)/2} + \cdots = d_\kappa \log f_n(\zeta) + \frac{d_\lambda}{2-\lambda} f_n(\zeta)^{-(\lambda-\kappa)} + \cdots + \text{Const}.$$

For $\zeta = \alpha$ we have $f_n(\zeta) = \infty$; this is compatible with (64) and (64') only if $\kappa > 2$, since only in this case does the right side remain finite. Hence

$$(62') \quad z\chi(\zeta) = c_\kappa z^{-(\kappa-1)} + c_\lambda z^{-(\lambda-1)} + \cdots \quad \text{with} \quad \kappa - 1 > 1.$$

Thus $\int z\chi(\zeta)dz$ remains finite at $z=\infty$ and $\int\{z\chi(\zeta)-L_n(z)\}dz$ is finite everywhere on $\mathfrak{R}_n$ and possesses a uniform derivative. The expression is therefore an integral of the first kind $w(z)$. Introducing the values of $\chi(\zeta)$ and $L_n(z)$, we finally obtain the following differential equation for $f_n(\zeta)$:

$$(65)\quad \frac{f_n(\zeta)}{\zeta^2 f'_n(\zeta)^2}\left(\frac{1}{\zeta^{n-1}}+\frac{2a_2}{\zeta^{n-2}}+\cdots+(n-1)a_n+\cdots+2\bar{a}_2\zeta^{n-2}+\zeta^{n-1}\right)$$
$$+\sum_{l=2}^{n}\frac{a_{nl}'}{l-1}\tau'_{l-1}(f^n(\zeta);f_n(0))=w'(f_n(\zeta)).$$

From this differential equation it follows that $\mathfrak{R}_n$ has at most $(n-1)$ branch points in the interior of D_n which can only be situated at those $(n-1)$ values ζ in $|\zeta|<1$, for which $\frac{1}{\zeta^{n-1}}+\frac{2a_2}{\zeta^{n-2}}+\cdots+2\bar{a}_2\zeta^{n-2}+\zeta^{n-1}=0$. For suppose that a branch point z_b corresponds to a point ζ_b, which is not one of the mentioned roots; then $\frac{1}{\zeta^2 f'_n(\zeta)^2}[\cdots]$ becomes infinite as $(\zeta-\zeta_b)^{-2(k-1)}$, while the other terms of (65) become infinite as $(\zeta-\zeta_b)^{-(k-1)}$ at most. Thus, the terms cannot cancel and (65) is impossible.

It is further easy to show that there are no points on $\mathfrak{R}_n$ exterior to D_n and that D_n is, therefore, a slit domain bounded by analytic curves. For let t be a point on $\mathfrak{R}_n$ exterior to D_n; choose $z_0\cdots z_\gamma$ in the neighborhood of t and exterior to D_n. The above analysis is once more applicable, only $\chi(\zeta_0),\cdots\chi(\zeta_\gamma)$ must be omitted in all formulas. Instead of (58) we get the formula

$$(58')\quad a\sum_{i=\gamma+1}^{\gamma+p}\chi(\zeta_i)[\tau(z_i;z)-\tau(y;z)]+\left\{\frac{f_n(\xi)}{f_n(\xi)-a}[\tau(f_n(\xi);z)-\tau(y;z)]\right\}_{n\text{-th}}$$
$$+\frac{a_n}{a}[\tau(f_n(0);z)-\tau(y;z)]=\sum_{k=1}^{\gamma}\lambda_k\frac{dw_k(z)}{dz}$$

which holds for each z in the neighborhood of t and for arbitrary a. Therefore, according to the principle of permanence, it must hold everywhere. But in the neighborhood of $\zeta=0$, the right side of the equation remains finite, while the left becomes infinite, which shows the contradiction.

EINSTEIN INSTITUTE OF THE HEBREW UNIVERSITY,
JERUSALEM, PALESTINE.